Sorana Săveanu

Prelucrarea și analiza datelor în științele sociale. Utilizarea programului Excel

Presa Universitară Clujeană

Sorana Săveanu

Prelucrarea și analiza datelor în științele sociale. Utilizarea programului Excel

Presa Universitară Clujeană

2020

Referenți științifici

Prof. univ. dr. habil. Sergiu Bălțătescu

Prof. univ. dr. habil. Adrian Hatos

ISBN 978-606-37-0822-0

© Autoarea volumului. Toate drepturile rezervate. Reproducerea integrală sau parțială a textului, prin orice mijloace, fără acordul autoarei, este interzisă și se pedepsește conform legii.

Universitatea Babeș-Bolyai
Presa Universitară Clujeană
Director: Codruța Săcelean
Str. Hasdeu nr. 51
400371 Cluj-Napoca, România
Tel/Fax: (+40)-264-597.401
E-mail: editura@editura.ubbcluj.ro
http://www.editura.ubbcluj.ro/

CUPRINS

ASPECTE GENERALE PRIVIND UTILIZAREA PROGRAMULUI EXCEL PENTRU CALCUL TABELAR 9
Orientarea în pagină şi formatarea paginilor de lucru 9
Inserarea, ştergerea şi modificarea dimensiunii rândurilor şi a coloanelor 11
Blocarea rândurilor şi a coloanelor pe ecran 12
Multiplicarea foilor de calcul 13
Formatarea celulelor 15
Formatarea tabelelor 27
Formatarea condiționată 29
Sortarea tabelelor 33
Filtrarea tabelelor 34
Eliminarea duplicatelor 37
Validarea datelor 38
FORMULE ŞI FUNCȚII. ASPECTE INTRODUCTIVE 43
Adăugarea formulelor și funcțiilor 43
Elementele formulelor și funcțiilor 44
Operatori matematici utilizați în cadrul formulelor 45
Opțiunea auto-umplere 45
Utilizarea referințelor 47
Introducerea funcțiilor 50
FUNCȚIILE MATEMATICE ŞI STATISTICE: SUM, PRODUCT, SUMPRODUCT, AVERAGE, COUNT, COUNTA, COUNTBLANCK, MIN, MAX, SUBTOTAL 55
Realizarea Subtotalurilor accesând meniul DATA>Subtotal 68
FUNCȚIILE CONDIȚIONALE: COUNTIF, COUNTIFS, SUMIF, SUMIFS, AVERAGEIF, AVERAGEIFS 71
FUNCȚIILE LOGICE: AND, OR, IF, NOT 87
Prezentarea funcțiilor logice 87
Utilizarea funcțiilor logice 89

STATISTICA DESCRIPTIVĂ A DATELOR: FREQUENCY, MODE, MEDIAN, QUARTILE, SQRT, VAR, STDEV, CONFIDENCE.NORM 95
Realizarea tabelelor de frecvență..... 95
Tabelele de frecvență pentru variabilele calitative 95
Tabelele de frecvență pentru variabilele cantitative (numerice)..... 99
Calculul principalilor indicatori statistici 103
Calcularea principalilor indicatori statistici pe baza tabelului de frecvență . 104
Calcularea principalilor indicatori statistici pe baza valorilor variabilei folosind funcțiile statistice..... 110
ANALIZA STATISTICĂ BIVARIATĂ. TESTAREA RELAȚIEI DE ASOCIERE ÎNTRE VARIABILELE CATEGORIALE: CHISQ.TEST, CHISQ.INV.RT, CHISQ.DIST.RT 117
Realizarea tabelelor de contingență (asociere) 117
Testul chi-square (testul chi-pătrat)..... 120
REPREZENTAREA GRAFICĂ A DATELOR.....129
Care sunt elementele principale ale unui grafic?..... 130
Cum realizăm un grafic în Excel? 131
Ce facem în situația în care graficul realizat nu cuprinde și valorile variabilelor?..... 134
Cum alegem tipul de grafic?..... 136
Cum ne asigurăm că graficul realizat cuprinde toate informațiile? 139
Cum adăugăm informațiile pentru legendă, dacă acestea lipsesc? 142
Cum adăugăm valorile etichetelor? 144
Cum modificăm design-ul graficului?..... 145
Cum modificăm un grafic?..... 146
Cum mutăm un grafic? 147
Cum realizăm formatări suplimentare ale graficului? 147
Alte modalități de prezentare vizuală a datelor 149
Câteva recomandări și observații finale privind reprezentarea grafică..... 151

UTILIZAREA FUNCȚIILOR PENTRU GESTIUNEA INFORMAȚIILOR: CONCATENATE, PROPER, MID, LEFT, RIGHT, LEN, UPPER, LOWER, ISEVEN, ISODD, ISNUMBER, ISTEXT 153
Funcțiile pentru text..... 153
Funcțiile pentru informații..... 158
Resurse bibliografice recomandate.....161

Prefață

Acest *Handbook* se adresează în special studenților din domeniul științelor socio-umane, care sunt preocupați de activitatea de cercetare și care doresc să își îmbunătățească competențele privind prelucrarea și analiza datelor folosind programul Excel. Totodată, aplicațiile prezentate în carte sunt utile tuturor celor care utilizează datele în activitatea profesională curentă și urmăresc să devină cât mai operativi în managementul și prelucrarea datelor și informațiilor.

Tuturor celor care ajung să parcurgă informațiile cuprinse în această carte le recomand să nu se rezume la această lectură, ci să realizeze individual aplicațiile prezentate în capitole. Este important să înțelegem că doar utilizând programul Excel în mod uzual și prin familiarizarea cu diferitele funcții și formule, vom reuși cu adevărat să profităm de avantajele acestui program pentru calcul tabelar.

Prima parte a cărții detaliază aspecte generale privind utilizarea programului Excel și aspecte introductive privind elaborarea funcțiilor și a formulelor. Cea de-a doua parte de cărții cuprinde capitole în care sunt prezentate principalele funcții și formule folosite în prelucrarea și analiza datelor: funcțiile matematice și statistice, funcțiile condiționale, funcțiile logice, formule și funcții pentru analiza statistică descriptivă, funcțiile text și informaționale. Un capitol de adresează reprezentării grafice a datelor.

În fiecare capitol sunt incluse exemple și aplicații ale informațiilor prezentate. Cele mai multe dintre datele prezentate sunt fictive și sunt utilizate doar în scop didactic, aplicativ, astfel ele nu pot fi folosite în argumentările științifice. În cazul în care baza de date prezentată are o sursă reală (este vorba despre date de la Institutul Național de Statistică), această sursă a datelor este menționată direct în text.

În ceea ce privește capitolele privind analiza statistică a datelor, acestea nu își propun introducerea noțiunilor generale necesare acestui tip de analiză. Capitolele se concentrează pe prezentarea modalităților de calcul a principalilor indicatori statistici și utilizarea principalelor formule și funcții pentru statistica descriptivă. În acest sens, este recomandat ca cititorii să dețină în prealabil cunoștințe minime de statistică.

În carte sunt prezentate doar câteva dintre funcțiile și formulele utilizate frecvent în managementul, prelucrarea și analiza datelor. Sunt o serie de alte funcții disponibile în Excel care de asemenea sunt extrem de utile pentru analiza datelor, respectiv sunt o serie de caracteristici care nu au fost incluse în această carte, iar acest lucru nu înseamnă că acestea ar fi mai puțin importante. În analiza datelor, este recomandată de asemenea utilizarea tabelelor pivot, respectiv a panourilor de comandă, iar în analiza statistică sunt folosite și testele de corelație, testele pentru compararea mediilor, regresia etc.. Toate acestea nu sunt incluse în acest *Handbook*, dar autoarea își propune să adauge aceste aspecte într-un alt material suport.

Despre autoare

Sorana Săveanu este lector dr. la Facultatea de Științe Socio-Umane din cadrul Universității din Oradea. În perioada 2003-2016 și-a desfășurat activitatea ca cercetător științific, fiind implicată în elaborarea și implementarea proiectelor de cercetare derulate de către Departamentul de Sociologie și Asistență Socială. Începând din 2016, alături de activitatea de cercetare științifică, desfășoară activități didactice având discipline precum *Prelucrarea informatizată a datelor sociale, Educație și societate, Managementul sistemelor informaționale, Managementul proiectelor sociale*. Rezultatele activităților de cercetare sunt prezentate în peste 20 de articole publicate în reviste de specialitate, 18 capitole și studii în volume colective și peste 30 de lucrări prezentate la conferințe internaționale.

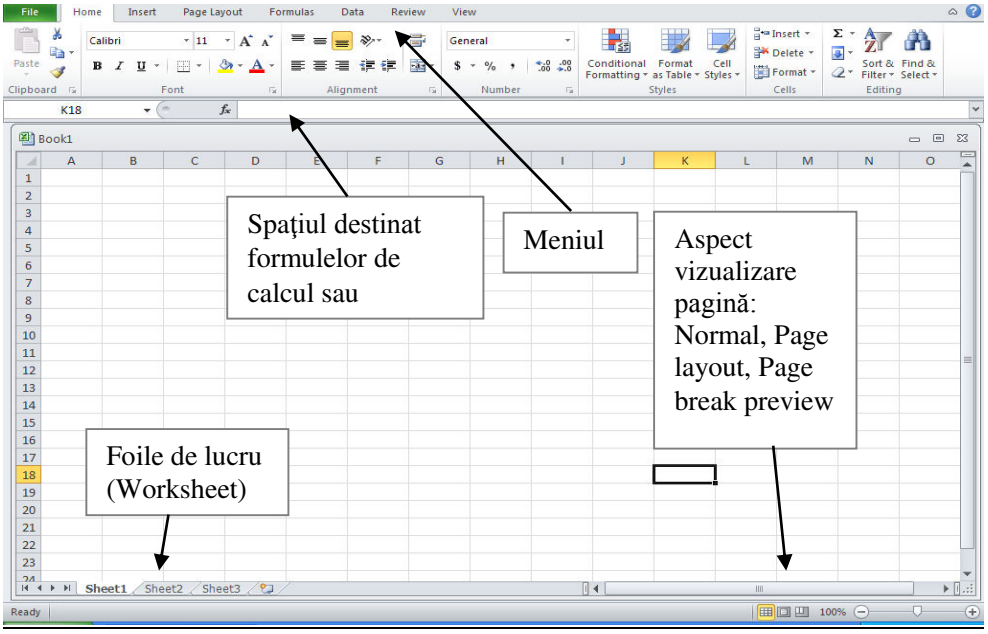
1 ASPECTE GENERALE PRIVIND UTILIZAREA PROGRAMULUI EXCEL PENTRU CALCUL TABELAR

Orientarea în pagină și formatarea paginilor de lucru

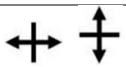
Microsoft Excel este un program de calcul tabelar. Foaia de calcul este organizată pe coloane și rânduri. Un fișier Excel (Registru de calcul) cuprinde mai multe foi de calcul (*Worksheet*). Acestea (în număr de 1 sau 3 la deschiderea unui fișier nou) au o etichetă generică (*Sheet1, Sheet2, Sheet3*) și le putem vedea la baza ferestrei. Operațiile se pot realiza pe o singură foaie de calcul, dar sunt permise și legături între foile de calcul din același fișier și chiar legături cu foi de calcul din alte fișiere.

Într-o foaie de calcul ne vom orienta pornind de la capul de coloană (coloanele sunt marcate cu litere A, B, C etc.) și capul de rând (rândurile sunt notate cu cifre 1, 2, 3 etc.). O celulă se va afla la intersecția dintre o coloană și un rând.

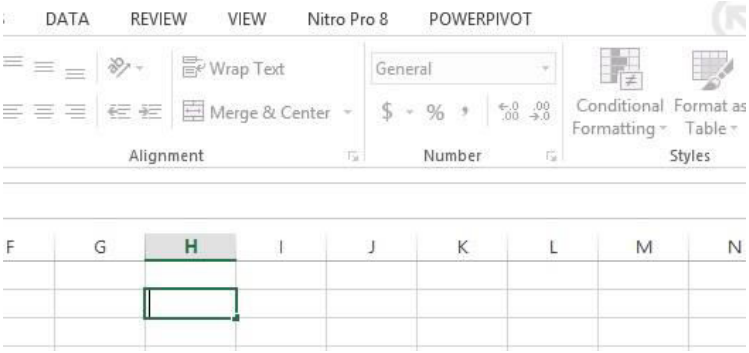
De exemplu: Celula K18 se află la intersecția dintre coloana K și rândul 18. Celula selectată va fi indicată cu un chenar îngroșat, iar capul de rând și cel de coloană vor fi marcate cu portocaliu.


Spaţiul destinat formulelor de calcul sau
Meniul
Aspect vizualizare pagină: Normal, Page layout, Page break preview
Foile de lucru (Worksheet)

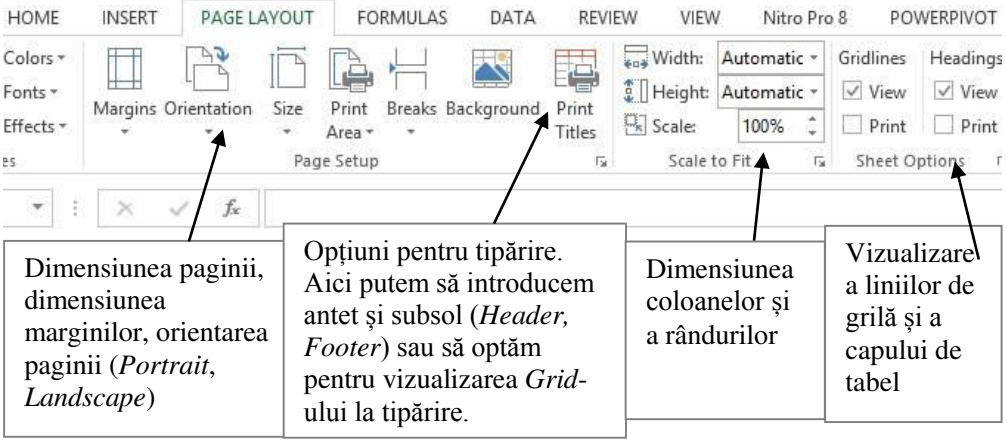
În funcție de modul în care ne orientăm în pagină, cursorul are diferite aspecte:

	= pentru selecţia celulelor
	= mutarea celulei/celulelor selectate (*Drag&Drop*)
	= multiplicarea informaţiei (*Auto Fill*)
	= selecția coloanelor sau a rândurilor
	= modificarea dimensiunii coloanelor sau rândurilor

În cazul în care apăsăm dublu-click pe o celulă, cursorul se transformă în Editorul de text (I). Acest aspect al cursorului ne indică faptul că ne aflăm în interiorul unei celule și din acest motiv cea mai mare parte din meniu este inactivă. Pentru a ieși din editorul de text, fie apăsăm tasta ESC, fie selectăm cu mouse-ul o altă celulă.

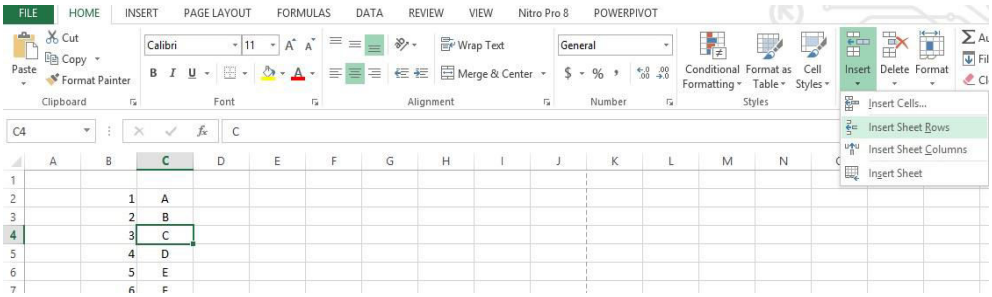


Meniul **PAGE LAYOUT** conține opțiunile privind modificarea aspectului paginii de lucru.



Inserarea, ștergerea și modificarea dimensiunii rândurilor și a coloanelor

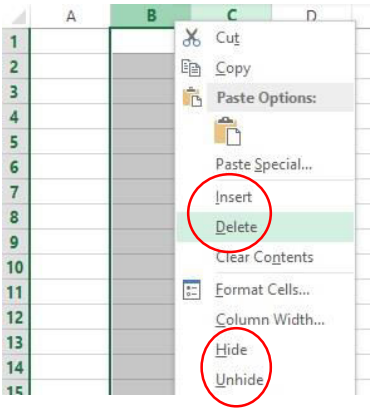
Numărul de rânduri şi coloane, respectiv dimensiunea acestora, sunt prestabilite pentru o foaie de calcul nouă. Putem să inserăm sau să ștergem una sau mai multe coloane ori rânduri accesând meniul **HOME>Cells.** Pentru inserare, respectiv ștergere, este importantă selecția anterioară. În cazul inserării unui rând, **> Insert Rows** va insera un nou rând deasupra celui în care avem selecția (fie selectăm doar o celulă, fie un rând întreg). În imaginea de mai jos, accesând **> Insert Sheet Rows** vom adăuga un rând nou între rândul 3 și 4, deoarece selecția este pe rândul 4.



La fel, în cazul inserării unei noi coloane, inserția se va face la stanga coloanei pe care se află selecția.

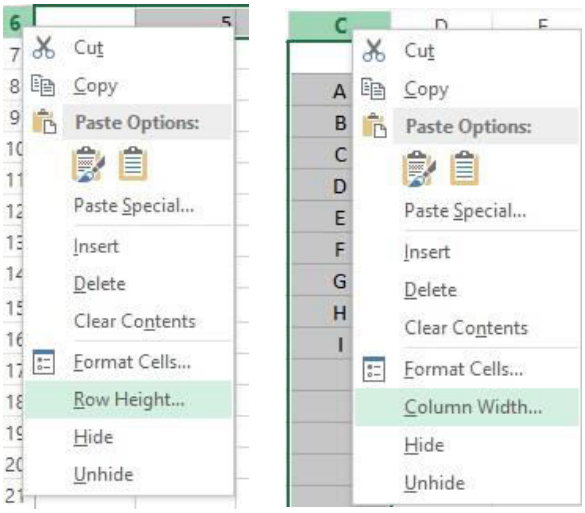
În ceea ce privește ștergerea coloanelor, respectiv a rândurilor, putem să selectăm fie un rând sau o coloană, fie mai multe şi accesăm **>Delete Sheet Rows**, respectiv **>Delete Sheet Rows.** Apăsarea tastei **Delete** nu șterge un rând sau o coloană, ci doar informația cuprinsă în celula sau celulele selectate.

Același lucru îl putem face fără să mai accesăm meniul **HOME**, ci direct apăsând click-dreapta pe capul de tabel (unde coloanele sunt notate cu A, B, C...) **> Insert** sau **Delete,** pentru coloane, respectiv pentru rânduri – click-dreapta în stânga (unde rândurile sunt numerotate 1, 2, 3...) **> Insert** sau **Delete.**



Putem să ascundem coloanele, respectiv rândurile unui tabel, fără să le ștergem complet: click-dreapta pe coloana/rândul (sau coloanele/rândurile) pe care doriți să îl (să le) ascundeți **Hide**. Dacă doriți ca rândul/coloana ascuns/ă anterior să se vadă din nou: **Unhide**

Dimensiunea rândurilor, respectiv a coloanelor, poate fi modificată apăsând click-dreapta **>Row Height** (pentru rânduri), respectiv **>Column Widh** (pentru coloane).

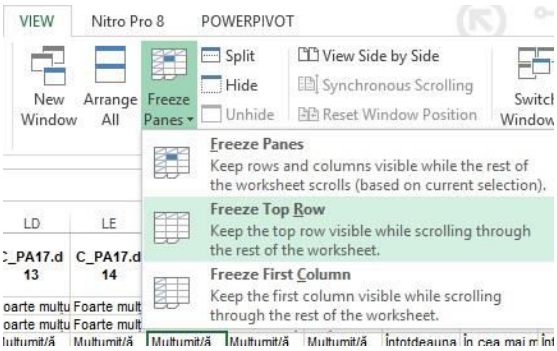


Dimensiunea rândurilor, respectiv a coloanelor, poate fi modificată și manual deplasând mouse-ul pe linia ce desparte două rânduri, respectiv coloane. Ținând apăsat butonul mouse-ului, trageți linia până ajunge la dimensiunea dorită. Pentru a putea face aceste modificări, cursorul trebuie să arate astfel ↕ (modificare dimensiune rânduri) sau ↔ (modificare dimensiune coloane).

Blocarea rândurilor și a coloanelor pe ecran

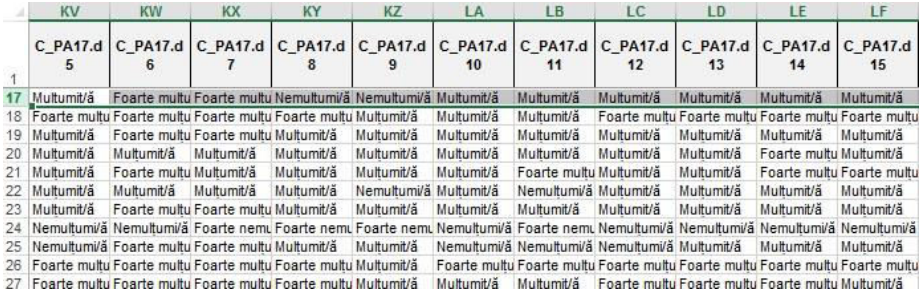
De cele mai multe ori în Excel avem de a face cu baze de date mari, cu multe intrări, atât pe coloane, cât și pe rânduri. Introducerea datelor, respectiv urmărirea informației într-o foaie de calcul poate să ridice probleme în acest sens. Dacă avem multe rânduri, spre exemplu, pe măsură ce vom derula în jos foaia de calcul, primul rând (care de regulă cuprinde capul de tabel cu etichetele variabilelor) nu mai este vizibil. Putem opta în această situație pentru blocarea rândurilor, respectiv a coloanelor.

Pentru blocarea (înghețarea – *Freeze*) rândurilor este necesar să selectăm rândul de sub cel pe care dorim să îl blocăm. Pentru blocarea coloanelor este necesar să selectăm coloana aflată în dreapta celei pe care dorim să o blocăm. După selecție, accesăm din meniul **VIEW>Freeze Panes**. Aici putem opta pentru blocarea primului rând sau a primei coloane, fie vom bloca un anumit rând sau o anumită coloană în funcție de cum am făcut anterior selecția.



De exemplu: O bază de date unde nu am blocat capul de tabel va fi dificil de gestionat deoarece nu ne putem da seama care sunt variabilele.

Dacă în schimb vom bloca primul rând, vom vedea codurile variabilelor. Se poate observa că rândul 1 este blocat, iar restul ecranului începe de la rândul 17.

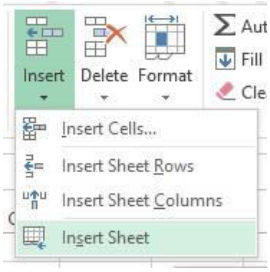


Pentru a debloca rândurile, respectiv coloanele, vom folosi tot din meniul **VIEW>Unfreeze Panes,** care este activ după ce inițial am blocat rândurile sau coloanele.

Multiplicarea foilor de calcul

La inițierea unui fișier nou avem una sau trei foi de calcul disponibile (în funcție de versiunea programului). Pe măsură ce lucrăm și în funcție de modul în care dorim să gestionam datele incluse în documentul nostru, putem să inserăm noi foi de calcul, să le multiplicăm sau să le mutăm pe cele existente. Putem să facem acest lucru în mai multe feluri.

Inserarea unei noi foi de calcul o realizăm accesând meniul **HOME>Insert Sheet**. Pagina nouă este inserată înaintea celei selectate.

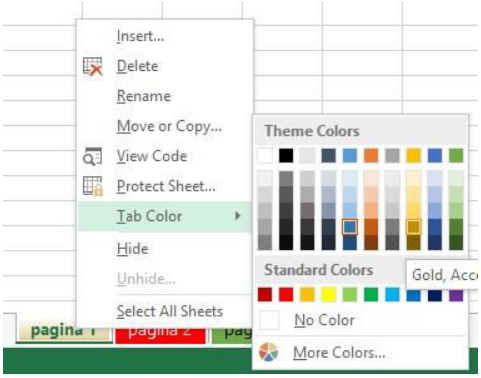


Numele paginilor poate fi schimbat cu dublu-click pe titlu, intrând astfel în editare de text sau aplicând click-dreapta **>Rename**. Am denumit foile de calcul pagina 1, 2 și 3.

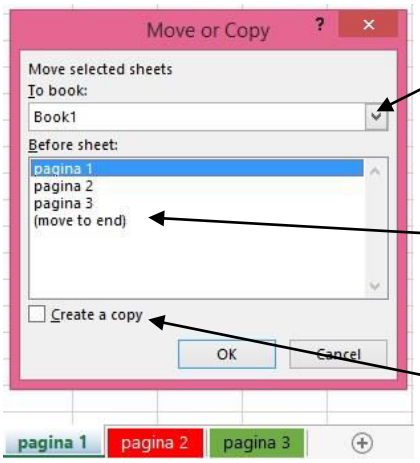
Putem să inserăm noi pagini mai simplu: doar apăsând butonul sau aplicând click-dreapta pe numele foii de calcul **>Insert**. Pentru ștergerea unei foi de calcul accesăm **>Delete.**

Poziția unei foi de calcul într-un registru poate fi modificată doar mutând cu mouse-ul foaia în locul în care dorim (*Drag & Drop*).

Este bine atunci când lucrăm cu registre mari, cu multe foi de calcul, să gestionăm în așa fel documentul încât să nu fim puși în dificultate când căutăm o anumită informație. În acest sens, este recomandat să dăm titluri sugestive foilor de calcul, iar în plus, putem să folosim și culori. Adăugarea unei culori o facem aplicând click-dreapta pe numele foii de calcul **>Tab Color**, iar aici alegem culoarea pe care o dorim.



Foile de calcul pot fi mutate sau copiate, fie în acelaşi fişier, fie într-un fişier diferit. Realizăm acest lucru aplicând click-dreapta pe numele foii de calcul pe care dorim să o copiem sau să o mutăm **>Move or Copy**.



Fişierul (registrul de calcul) unde dorim să mutăm sau să copiem foaia.

Locul în cadrul unui fişier unde dorim să mutăm sau să copiem foaia.

Dacă dorim să multiplicăm (copiem) foaia de calcul, nu uităm să bifăm *Create a copy*.

Putem mai simplu să multiplicăm o foaie de calcul ţinând tasta CTRL apăsată, iar cu mouse-ul tragem (*Drag & Drop*) fila pe care dorim să o copiem.

Formatarea celulelor

Comenzile utilizate pentru modificarea aspectului şi a caracteristicilor celulelor se regăsesc în meniul **HOME.** Butoanele (instrumentele de lucru) sunt organizate pe grupuri logice.

Tab-ul **Font** se adresează formatării textului: tipul caracterelor, dimensiunea caracterelor, stiluri, culoare.

Tab-ul **Alignment** se referă la orientarea textului într-o celulă.

Din tab-ul **Number** selectăm tipul informaţiei din celule.

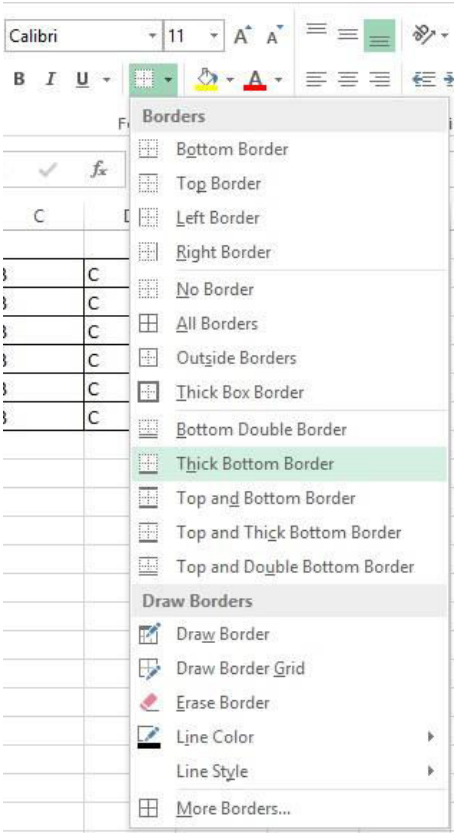
Chiar dacă sunt marcate ca vizibile liniile de grilă (*Gridlines*) şi pe ecran vedem în acest fel coloanele şi rândurile din foaia noastră de lucru, acest lucru nu înseamnă că am delimitat tabelul în sine pe care lucrăm.

De exemplu: Dacă am tipări tabelul din imaginea de mai jos din stânga, liniile coloanelor și ale rândurilor nu s-ar vedea. Acestea trebuie trasate așa cum sunt în imaginea din dreapta.

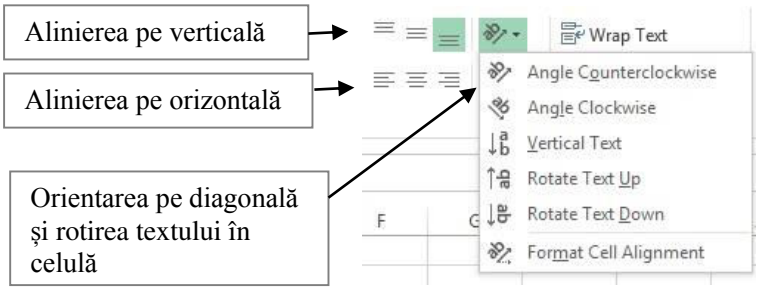
	A	B	C	D
1				
2		A	B	C
3		A	B	C
4		A	B	C
5		A	B	C
6		A	B	C
7		A	B	C
8				

	A	B	C	D
1				
2		A	B	C
3		A	B	C
4		A	B	C
5		A	B	C
6		A	B	C
7		A	B	C
8				

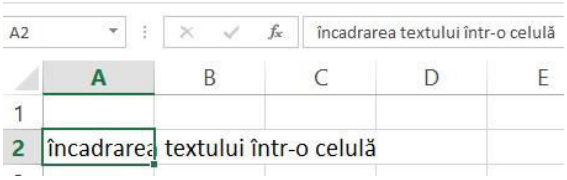
Vom trasa liniile pentru un tabel accesând meniul **HOME>Border**. Putem să alegem între diferite modalități de marcare a liniilor (toate liniile, doar liniile exterioare, doar liniile interioare etc.) și între diferite tipuri de linii (linie simplă, linie îngroșată, linie dublă etc.). Tot aici putem să alegem o altă culoare pentru linii.



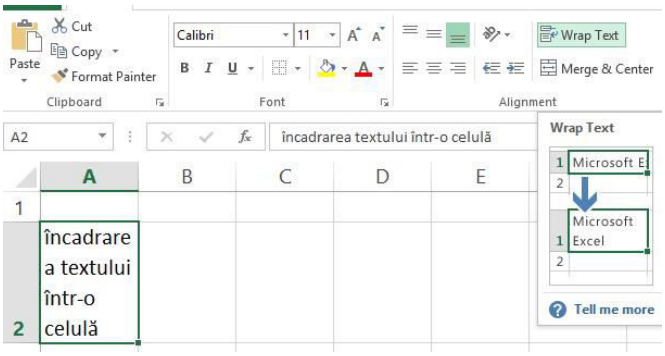
Pentru orientarea textului în celulă vom ține cont de faptul că fiind vorba de patru laturi ale unei celule, alinierea se va face atât pe orizontală (stânga, central, dreapta), pe verticală (sus, central, jos), cât şi pe diagonală.



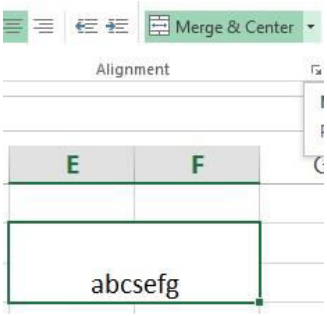
Foile de calcul au o dimensiune standard pentru coloane şi rânduri (pe care am văzut mai sus cum o putem modifica). Sunt multe situații în care informația introdusă într-o celulă nu se încadrează în celula respectivă şi se suprapune peste celulele din dreapta ei. În imaginea de mai jos avem o asemenea situație. Textul este prea mult ca să fie încadrat doar în celula în care a fost introdus (A2). Este o iluzie că textul "textului într-o celulă" se află în celulele B2, respectiv C2. Este suficient să ne uităm pe bara pentru funcții şi să observăm că selectând celula A2 tot textul se află introdus aici.



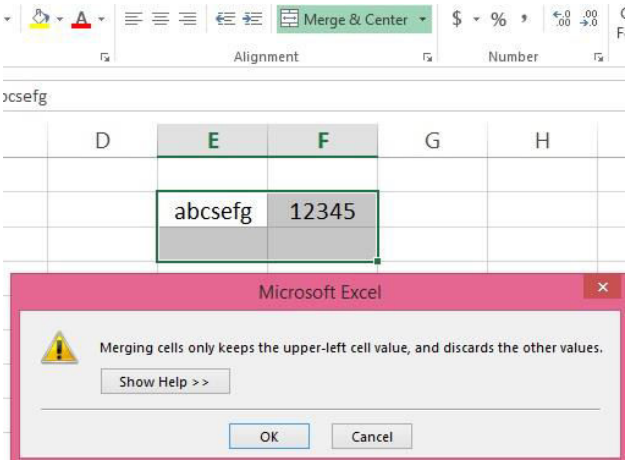
Pentru a nu se suprapune textul pe coloanele B şi C avem două posibilități: fie mărim dimensiunea coloanei, fie alegem opțiunea **Încadrare text**. Prima variantă este bună dacă tabelul nostru nu este unul mare, cu un număr mare de coloane, pentru care dacă am crește dimensiunea tuturor coloanelor nu vom mai putea urmări informația deoarece nu s-ar mai încadra pe ecran. Pentru a doua variantă vom accesa din meniul **HOME>Wrap Text**. Dacă butonul **Wrap Text** este activat, atunci informația va fi încadrată în celula în care este introdusă. Dacă dezactivăm acest buton, informația va reveni la situația în care se va suprapune peste celulele următoare din dreapta.



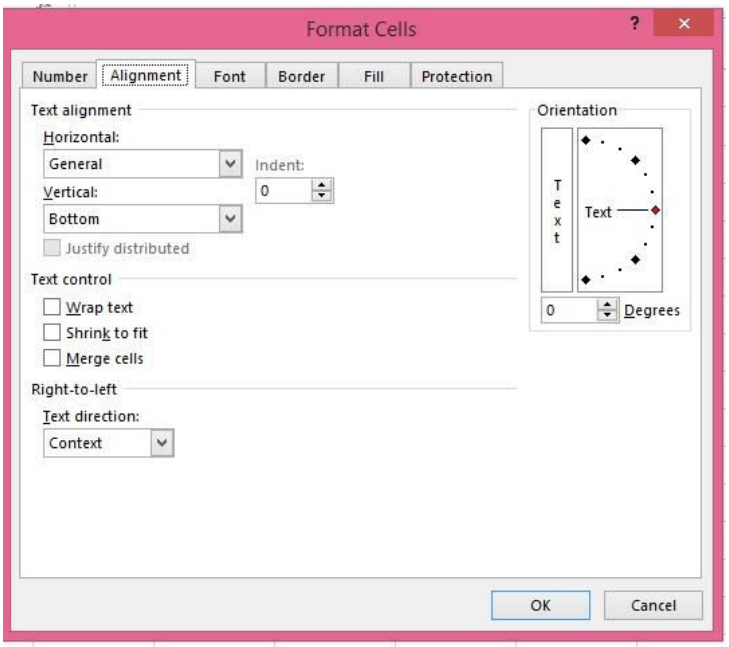
Într-o foaie de calcul putem opta pentru îmbinarea mai multor celule. Câte celule dorim să îmbinăm, atâtea vom selecta, apoi vom accesa meniul **HOME>Merge & Center**.



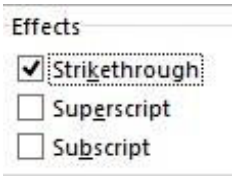
Această opțiune trebuie să o folosim cu atenție deoarece dacă toate celulele pe care dorim să le unim conțin informație, programul va păstra doar informația din prima celulă din stânga, iar restul va fi șters. De altfel, Excel ne avertizează de această situație printr-un mesaj în care ni se solicită să confirmăm că dorim să se șteargă informația respectivă.



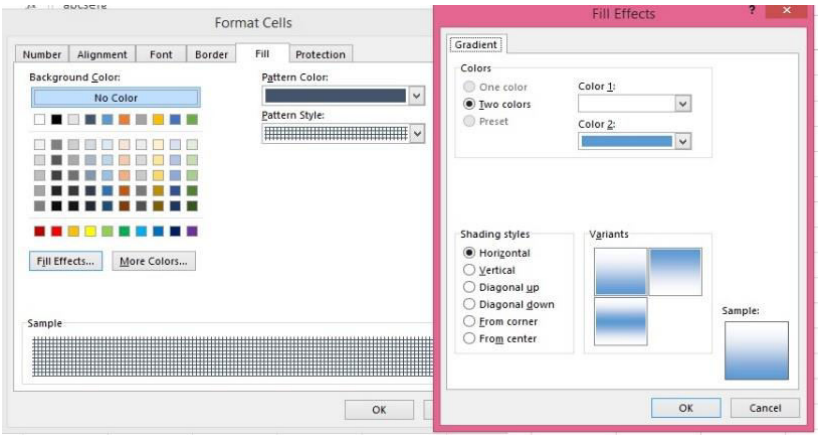
Toate opțiunile prezentate în această secțiune privind formatarea și orientarea informației într-o celulă le putem accesa deschizând fereastra **Format Cells** din meniul **HOME.** Fereastra se deschide dând click pe butonul din dreapta jos din tab-ul **Font**, **Alignament** sau **Number**. Fereastra **Format Cells** cuprinde următoarele file: Number (număr), Alignment (aliniament), Font (caracter), Border (bordură), Fill (umplere), Protection (protecție).



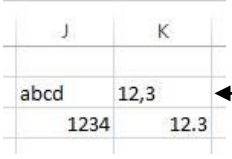
În această fereastră avem câteva opțiuni în plus. La *Alignament* avem opțiunea *Shrink to fit* care va micșora dimensiunea font-ului pentru a putea fi încadrat într-o celulă. La *Font* vom avea efecte de text: marcarea textului cu o linie (*Strikethrough* - ~~abc~~), indici și text exponențial (*Superscript* – a^2 și *Subscript* – a_2).



Tot în această fereastră putem să lucrăm cu efecte de umplere a celulelor. Totuși, este mai important să alocăm timp prelucrării și analizei datelor, decât formatării celulelor. Aceste aspecte privind formatarea sunt importante în măsura în care ne înlesnesc operațiunile ulterioare pe are dorim să le realizăm.

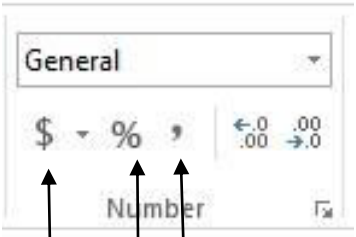


Tab-ul **Number** ne permite să selectăm tipul informației incluse în celule. Când introducem datele este selectat tipul de informație *General*. Acesta poate fi modificat, în funcție de tipul datelor cu care lucrăm. Totuși, înainte de a prezenta principalele tipuri, este important să menționăm diferența pe care o face Excel între caracterele tip *text* și caracterele tip *număr*. Dacă într-o celulă introducem text (litere) acestea se vor alinia automat către stânga. Dacă vom introduce numere (cifre), atunci informația se va alinia automat către dreapta. Acest aspect ne este de folos atunci când lucrăm de exemplu cu numere cu decimale.

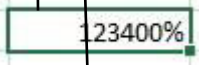


Informația "12,3" este citită de Excel ca și text deoarece în programul în care este realizat exemplul este setată virgula (,) ca separator pentru decimale.

Principalele tipuri de informații din celulă pe care le putem selecta: *Currency* (valută - introduce simboluri monetare specifice pentru diferitele tipuri de valută înaintea cifrei), *Percent* (procent - transformă formatul datelor tip procent), *Comma Style* (separator pentru mii, zeci de mii etc.). Dacă accesăm butonul din colțul din dreapta-jos, vom avea mai multe formate disponibile. Le vom discuta pe fiecare în parte.



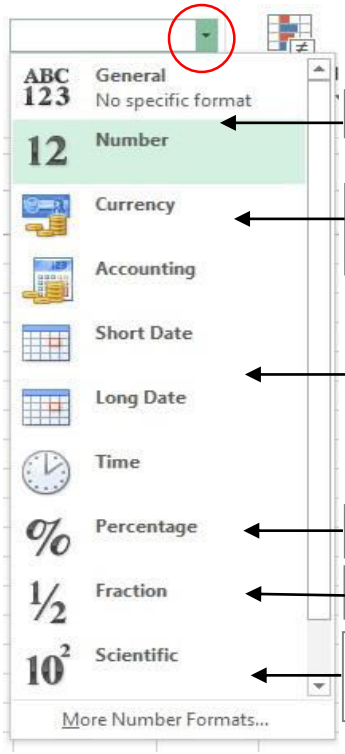
Formatul *Currency* adaugă semnul $ înaintea



Formatul *Percentage* adaugă semnul % după cifră și două 0-uri (*100). Mare atenție: acesta setare se referă doar la formatul informației. Nu calculăm în acest fel procentele pentru un tabel!

Formatul *Comma Style* utilizat pentru cifrele mari adaugă separatoare pentru mii, zeci de mii, sute de mii etc.

Alte tipuri de formate putem accesa apăsând săgeata care rulează întreaga listă:



Format pentru număr

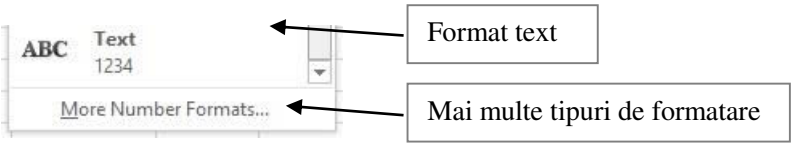
Formate pentru valori monetare

Formate pentru date calendaristice și timp

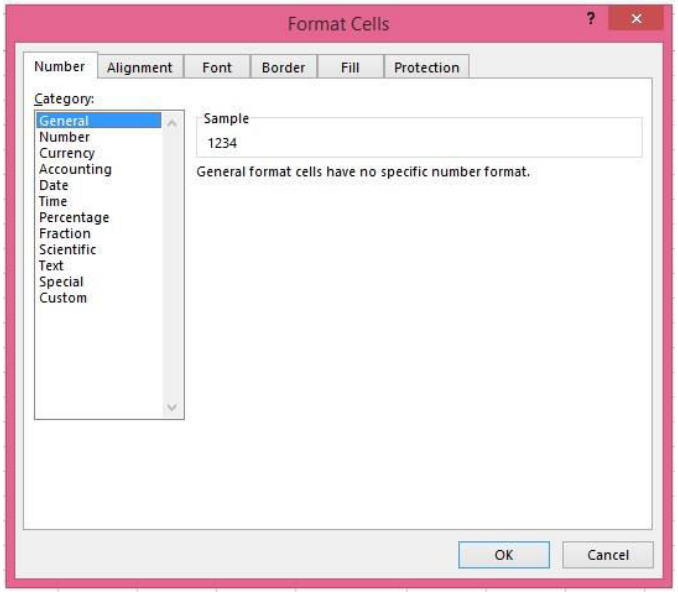
Format procent

Format fracție

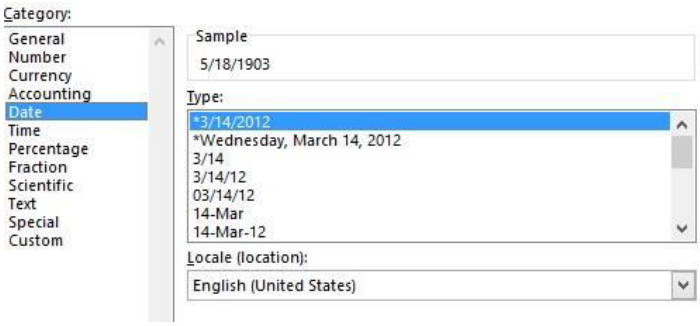
Format exponențial folosind E+*n* (E înmulțește numărul precedent cu 10 la puterea n)



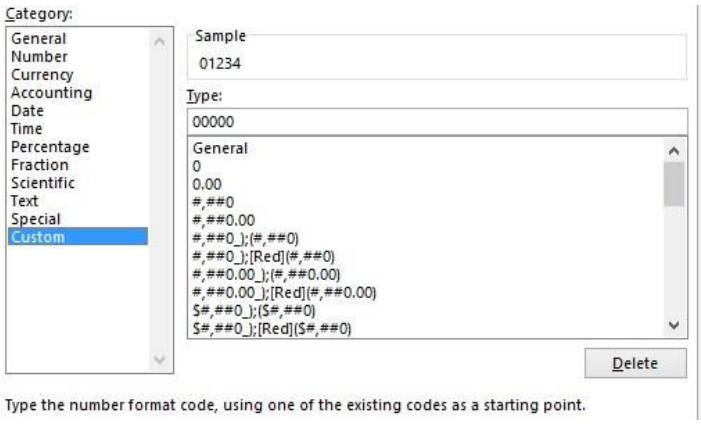
Dacă vom selecta **More Number Formats**… sau mai simplu vom selecta butonul din tab-ul **Number**, se va deschide următoarea fereastră de unde vom putea obține mai multe informații cu privire la diferitele tipuri de formatare:



Pentru formatarea *Date* (date calendaristice) avem disponibile mai multe tipuri de formate pentru care putem să stabilim și zona (formatul în funcție de locație):



Tot aici avem și formatul *Custom* care ne permite formatarea informației din celule în funcție de criterii specificate de noi sau putem să alegem direct un format disponibil în listă. Acest format ne permite să îmbunătățim modul de prezentare a datelor noastre.

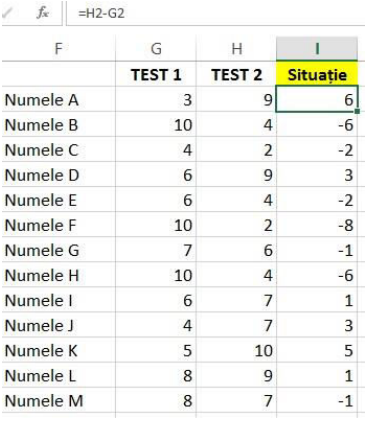


Custom Format ne permite să folosim anumite coduri, inclusiv culori, pentru diferitele formate de informații cu care lucrăm. În fereastra de mai sus adăugăm formatul pe care dorim să îl utilizăm pentru următoarele:

formatul utilizat pentru valorile negative; formatul utilizat pentru valorile pozitive; formatul utilizat pentru valorile 0; formatul utilizat pentru text

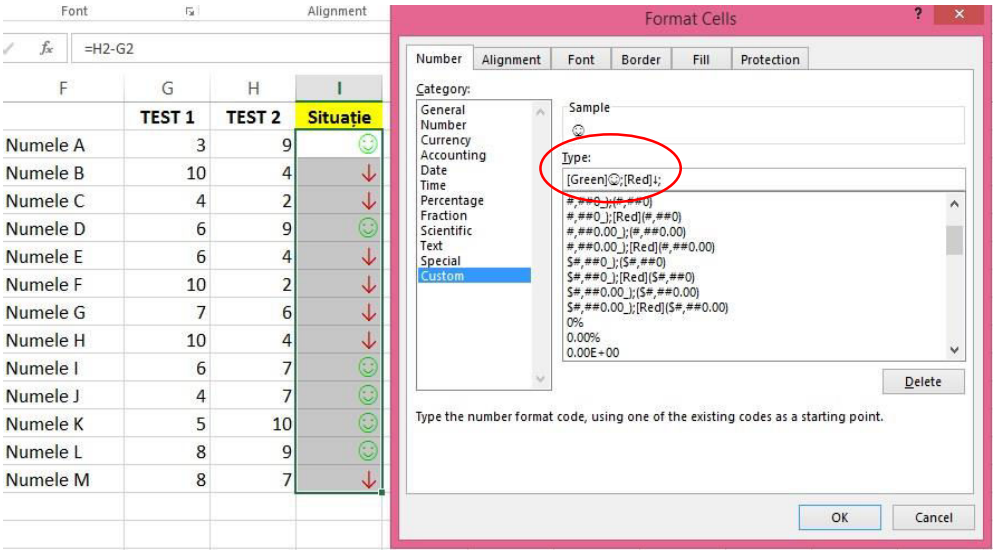
Dacă dorim să adăugăm o culoare, o vom scrie în paranteze drepte. De exemplu pentru culoarea roșie vom scrie [Red].

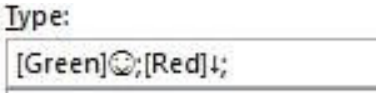
De exemplu: Am înregistrat rezultatele obținute la două teste date de către participanții la un curs. Dorim să marcăm situațiile în care participanții și-au îmbunătățit scorurile obținute, respectiv cei care au obținut un rezultat mai slab la a doua testare comparativ cu prima. În Excel vom face pur și simplu diferența dintre cele două note obținute. Dacă diferența indică o valoare pozitivă înseamnă că rezultatul la testarea a doua a fost mai bun decât la prima. Dacă în schimb diferența ne indică o valoare negativă, înseamnă că participanții respectivi nu și-au îmbunătățit rezultatele la a doua testare, înregistrând valori mai mici comparativ cu cele obținute la primul test.



F	G	H	I
	TEST 1	**TEST 2**	**Situație**
Numele A	3	9	6
Numele B	10	4	-6
Numele C	4	2	-2
Numele D	6	9	3
Numele E	6	4	-2
Numele F	10	2	-8
Numele G	7	6	-1
Numele H	10	4	-6
Numele I	6	7	1
Numele J	4	7	3
Numele K	5	10	5
Numele L	8	9	1
Numele M	8	7	-1

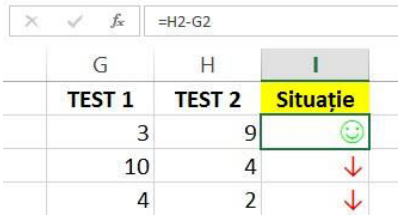
Putem să schimbăm modul în care vizualizăm situația fiecărui participant folosind *Custom Format*. Vom utiliza simbolul ☺ și culoarea verde pentru cei care au obținut rezultate mai bune la a doua testare, respectiv simbolul ↓ și culoarea roșie pentru cei care au obținut rezultate mai slabe.

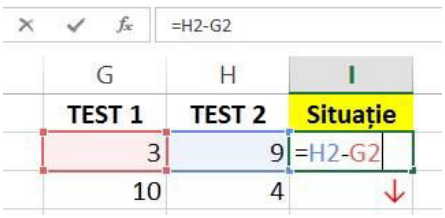




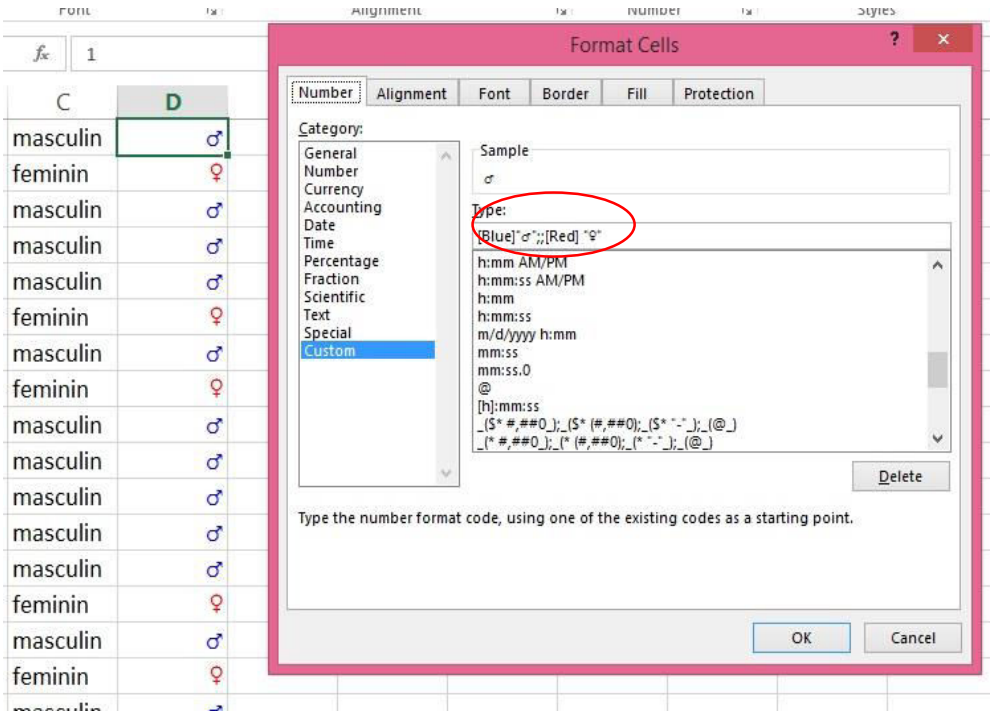
Vom scrie codul de mai sus care poate fi citit astfel: pentru valorile pozitive folosește culoarea verde și simbolul ☺, punct & virgulă (acesta face trecerea la următorul format), pentru valorile negative folosește culoarea roșie și simbolul ↓. Dacă am avea și valori 0 în baza noastră de date și am dori să avem un format și pentru acestea, atunci am adăuga punct & virgulă (;) și am insera codul pentru valori 0. Iar dacă în baza de date am avea și text și am dori să adăugăm un format și pentru acesta, atunci am adăuga iar punct & virgulă (;) și am insera codul pentru formatul text. Deoarece datele noastre nu conțin valori 0 și nici text, nu vom adăuga nimic în plus față de formatele pentru valorile negative și cele pozitive.

Se poate observa că indiferent de modul în care vizualizăm datele, în funcție de formatul ales, formula rămâne disponibilă "în interiorul" celulei. Dacă intrăm în editare-celulă (click-dublu pe celula selectată), atunci va fi disponibilă formula.





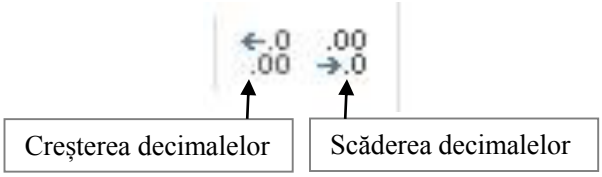
Un alt exemplu: Avem o bază de date în care am înregistrat sexul subiecților și am codificat această variabilă cu valorile 0 și 1 (0 pentru "feminin", 1 pentru "masculin"). Folosind *Custom Format* putem în locul codurilor 0 și 1 să vizualizăm simbolurile pentru sex.



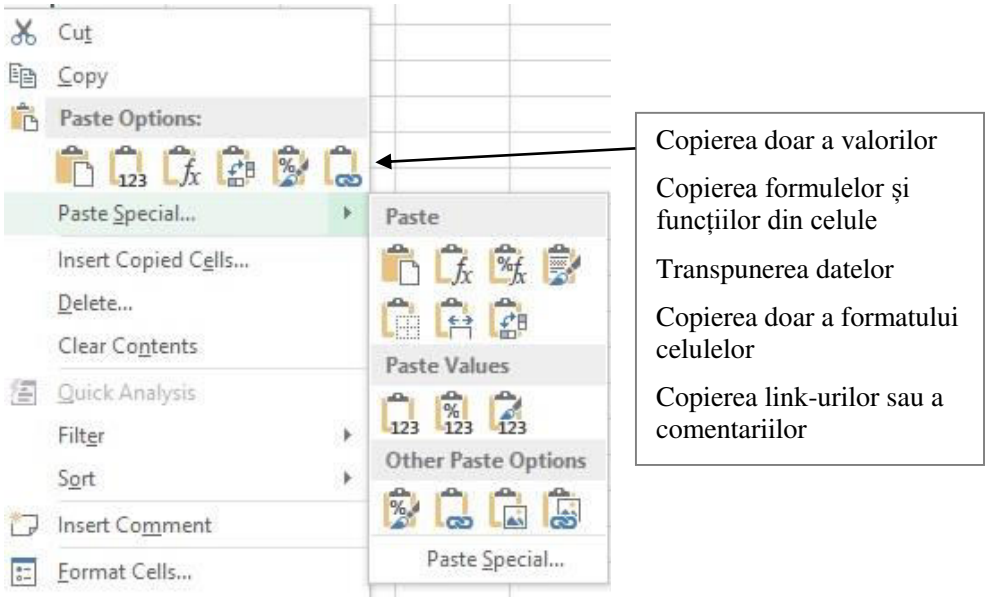
Codul utilizat este următorul:

Îl vom citi astfel: pentru valorile pozitive folosește culoarea albastră și simbolul ♂, punct & virgulă, pentru valorile negative nu vom adăuga niciun format (deoarece în datele noastre nu se regăsesc valori negative), astfel vom adăuga următorul punct & virgulă, pentru valorile 0 folosește culoarea roșie și simbolul ♀.

Creșterea, respectiv scăderea numărului de decimale o vom face folosind butoanele din meniul **HOME>Increase/Decrease Decimal**. Dacă lucrăm pe o selecție de celule, programul aplică numărul de decimale pornind de la prima celulă din selecție. De exemplu, dacă această primă celulă conține un număr cu două decimale, iar cifrele din toate celulele selectate au un număr diferit de decimale, apăsarea butonului **Increase decimal** (creșterea numărului de decimale) o dată va adăuga 3 decimale tuturor cifrelor din toate celulele selectate (deoarece numărul din prima celulă are două decimale).

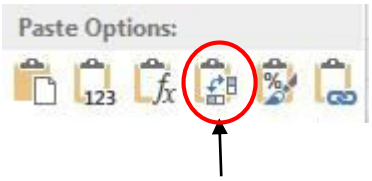


Pentru copierea și lipirea (**Copy & Paste**) celulelor avem mai multe variante din care putem să alegem în funcție de modul în care dorim să realizăm lipirea. Pentru a copia datele vom accesa **Copy** (sau tastăm **CTRL+C**), iar pentru lipire vom avea **Paste** (sau tastăm **CTRL+V**). Dacă dorim să accesăm diferitele variante de lipire vom da click-dreapta pe celula unde dorim să realizăm lipirea și vom alege din lista **Paste Special** varianta corespunzătoare formatului dorit:



De exemplu: Avem o bază de date în care am înregistrat valorile planificate pentru vânzarea unor produse pentru fiecare lună dintr-un an, respectiv valorile realizate. Am făcut diferența pentru a putea observa care sunt lunile în care nu am reușit să ne atingem valoarea planificată inițial.

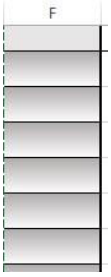
Dacă dorim să transpunem informația care este pe rânduri pe coloane și cea de pe coloane pe rânduri (capul de coloană să conțină lunile din anul de referință), vom folosi ***Paste Special>Transpose***:



	Jan-20	Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20
planificat	275	404	371	493	478	227	171	210	164	451	357	110
realizat	233	402	255	166	63	203	98	100	88	159	259	168
DIFEREN	-42	-2	-116	-327	-415	-24	-73	-110	-76	-292	-98	58

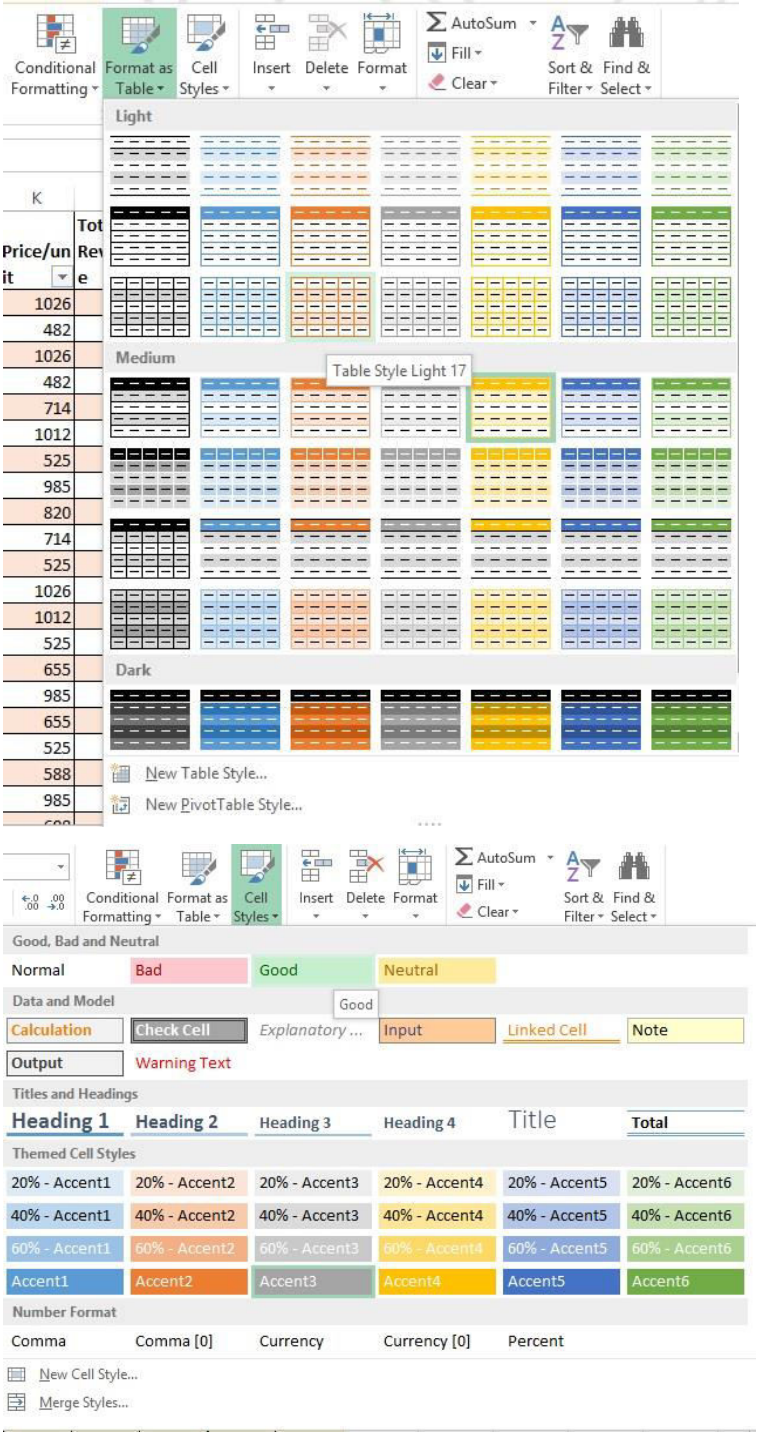
Dacă dorim să copiem din coloana a 3-a doar valorile și nu formulele, atunci vom folosi ***Paste Special>Values***:

Dacă dorim să copiem doar formatul celulelor din coloana a 3-a, fără valori sau formule, vom folosi ***Paste Special>Formats***:



Formatarea tabelelor

Programul Excel ne pune la dispoziție o serie de stiluri predefinite pentru formatarea tabelelor (***Format as Table***) și formatarea celulelor (***Cell Styles***).

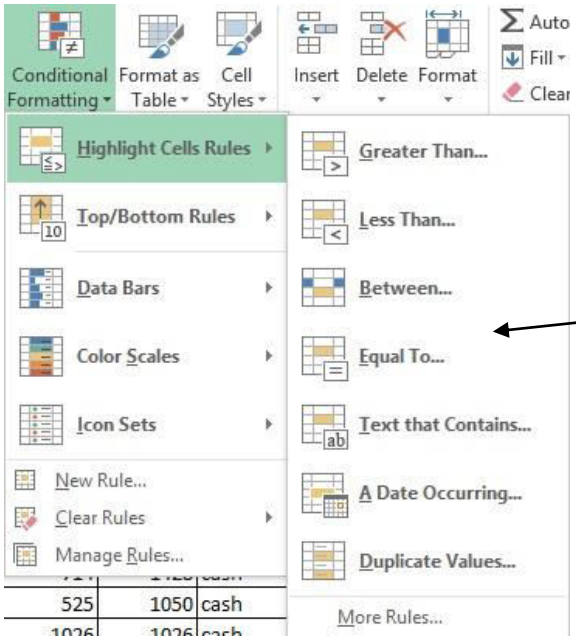

Conditional Formatting
Format as Table
Cell Styles
Insert
Delete
Format
AutoSum
Fill
Clear
Sort & Filter
Find & Select
Light
Medium
Dark
Table Style Light 17
New Table Style...
New PivotTable Style...
Good, Bad and Neutral
Normal
Bad
Good
Neutral
Data and Model
Calculation
Check Cell
Explanatory ...
Input
Linked Cell
Note
Output
Warning Text
Titles and Headings
Heading 1
Heading 2
Heading 3
Heading 4
Title
Total
Themed Cell Styles
Number Format
Comma
Comma [0]
Currency
Currency [0]
Percent
New Cell Style...
Merge Styles...

Formatarea tabelelor folosind ***Format as Table*** ne permite accesarea meniului **TABLE TOOLS>DESIGN** (care nu este activ dacă tabelul nu este formatat *As Table*). Pentru această opțiune putem să folosim și combinația de taste CRTL+T.

Formatarea tabelului în acest mod ne permite utilizarea anumitor instrumente, cum sunt filtrele (***Filter***), pivotarea (***Pivot Table***) și "feliatorul datelor" (***Slicer***).

Formatarea condiționată

În funcție de tipul datelor de care dispunem, putem să utilizăm **Conditional Formatting** (o formatarea condiționată) pentru a pune în evidență anumite aspecte ale datelor noastre. Putem să folosim următoarele moduri de evidențiere:



Evidențierea celulelor în funcție de anumite criterii:

Mai mare decât o anumită valoare

Mai mică decât o anumită valoare

Cuprinsă între două valori

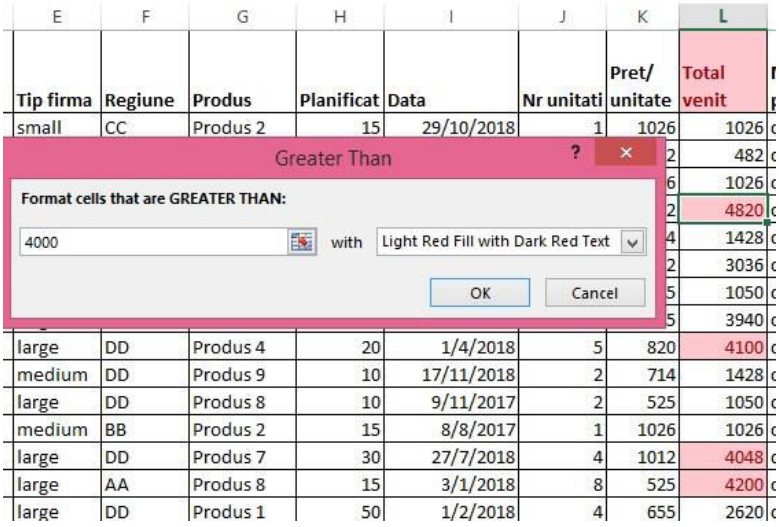
Valoare egală cu…

Text care conține…

O anumită dată

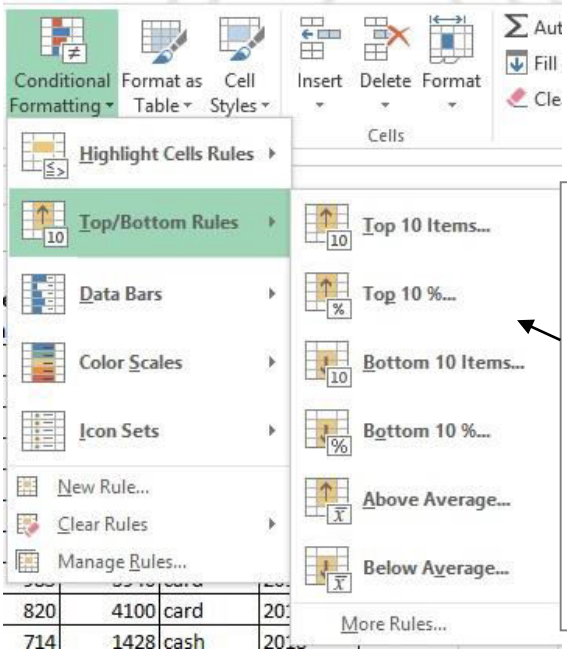
Valori duplicate

De exemplu: Avem o bază de date cu vânzările realizate într-o anumită perioadă de către angajați. Am înregistrat produsele vândute, numărul de unități vândute, prețul pe unitate și valoarea totală obținută. Putem să evidențiem de exemplu vânzările cu un venit de peste 4000 (***Highlight Cell Rules>Greater Than*** *4000*).



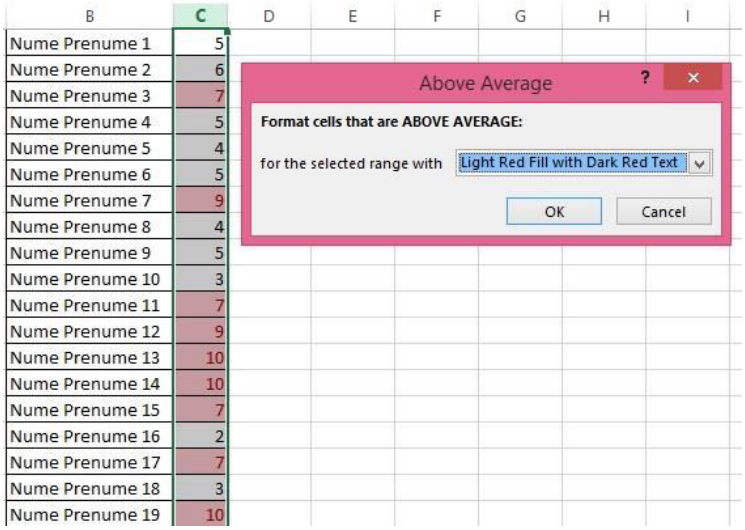
Tip firma	Regiune	Produs	Planificat	Data	Nr unitati	Pret/ unitate	Total venit
small	CC	Produs 2	15	29/10/2018	1	1026	1026
					2	482	
					6	1026	
					2	4820	
					4	1428	
					2	3036	
					5	1050	
					5	3940	
large	DD	Produs 4	20	1/4/2018	5	820	4100
medium	DD	Produs 9	10	17/11/2018	2	714	1428
large	DD	Produs 8	10	9/11/2017	2	525	1050
medium	BB	Produs 2	15	8/8/2017	1	1026	1026
large	DD	Produs 7	30	27/7/2018	4	1012	4048
large	AA	Produs 8	15	3/1/2018	8	525	4200
large	DD	Produs 1	50	1/2/2018	4	655	2620

O altă modalitate de evidențiere a informațiilor din celule utilizează ordonările:



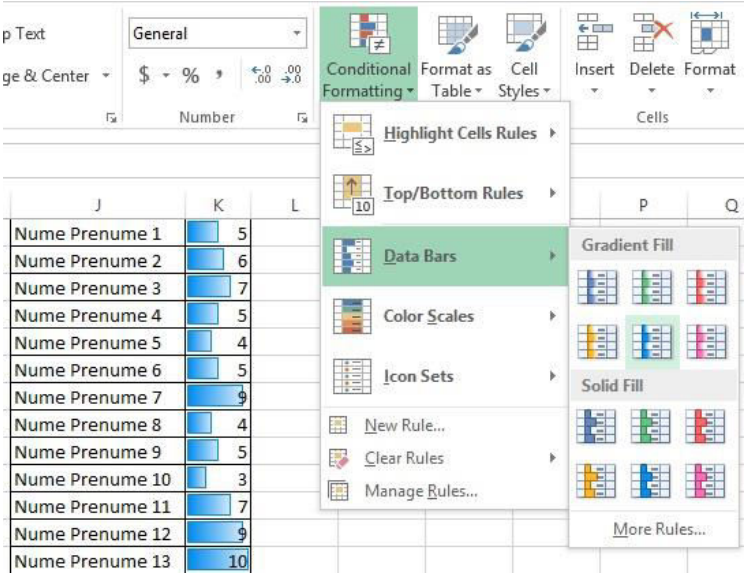
- Primele 10 cele mai mari valori
- Primele 10 cele mai mari valori exprimate în procente
- Primele 10 cele mai mici valori
- Primele 10 cele mai mici valori exprimate în procente
- Valori peste valoarea medie corespunzătoare șirului de date
- Valori sub valoarea medie corespunzătoare șirului de date

De exemplu: Dacă avem rezultatele obținute la un test de evaluare de către participanții la un curs, putem să evidențiem participanții care au obținut rezultate peste medie (***Top/Bottom Rules>Above Average***):



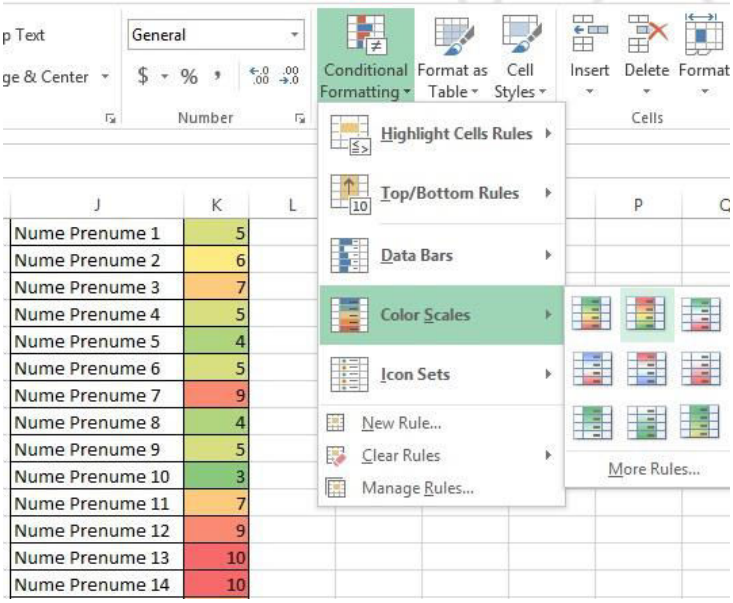
Nume Prenume 1	5
Nume Prenume 2	6
Nume Prenume 3	7
Nume Prenume 4	5
Nume Prenume 5	4
Nume Prenume 6	5
Nume Prenume 7	9
Nume Prenume 8	4
Nume Prenume 9	5
Nume Prenume 10	3
Nume Prenume 11	7
Nume Prenume 12	9
Nume Prenume 13	10
Nume Prenume 14	10
Nume Prenume 15	7
Nume Prenume 16	2
Nume Prenume 17	7
Nume Prenume 18	3
Nume Prenume 19	10

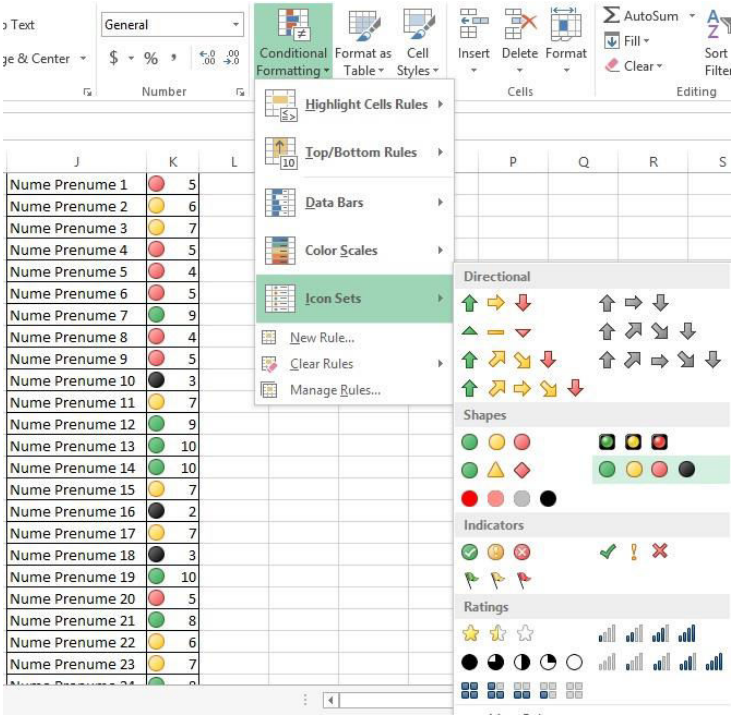
Tot cu o formatare condiționată putem să marcăm (*Data Bars*) fiecare celulă în funcție de valoarea pe care o conține. Dimensiunea barei va indica mărimea valorii incluse în celulă.



Nume Prenume 1	5
Nume Prenume 2	6
Nume Prenume 3	7
Nume Prenume 4	5
Nume Prenume 5	4
Nume Prenume 6	5
Nume Prenume 7	9
Nume Prenume 8	4
Nume Prenume 9	5
Nume Prenume 10	3
Nume Prenume 11	7
Nume Prenume 12	9
Nume Prenume 13	10

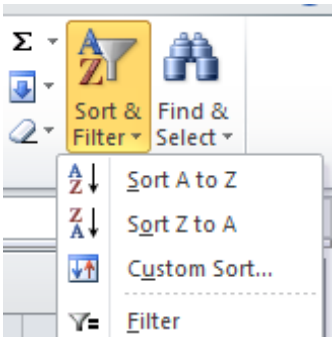
În același mod putem utiliza și diferite scale de culori (*Color Scales*) pentru a evidenția celulele în funcție de mărimea valorilor incluse sau diferite simboluri (*Icon Sets*) pentru fiecare celulă.


General
Conditional Formatting
Format as Table
Cell Styles
Insert
Delete
Format
Number
Cells
Highlight Cells Rules
Top/Bottom Rules
Data Bars
Color Scales
Icon Sets
New Rule...
Clear Rules
Manage Rules...
More Rules...
Nume Prenume 1 5
Nume Prenume 2 6
Nume Prenume 3 7
Nume Prenume 4 5
Nume Prenume 5 4
Nume Prenume 6 5
Nume Prenume 7 9
Nume Prenume 8 4
Nume Prenume 9 5
Nume Prenume 10 3
Nume Prenume 11 7
Nume Prenume 12 9
Nume Prenume 13 10
Nume Prenume 14 10


General
Conditional Formatting
Format as Table
Cell Styles
Insert
Delete
Format
AutoSum
Fill
Clear
Sort
Number
Cells
Editing
Highlight Cells Rules
Top/Bottom Rules
Data Bars
Color Scales
Icon Sets
New Rule...
Clear Rules
Manage Rules...
Directional
Shapes
Indicators
Ratings
Nume Prenume 1 5
Nume Prenume 2 6
Nume Prenume 3 7
Nume Prenume 4 5
Nume Prenume 5 4
Nume Prenume 6 5
Nume Prenume 7 9
Nume Prenume 8 4
Nume Prenume 9 5
Nume Prenume 10 3
Nume Prenume 11 7
Nume Prenume 12 9
Nume Prenume 13 10
Nume Prenume 14 10
Nume Prenume 15 7
Nume Prenume 16 2
Nume Prenume 17 7
Nume Prenume 18 3
Nume Prenume 19 10
Nume Prenume 20 5
Nume Prenume 21 8
Nume Prenume 22 6
Nume Prenume 23 7

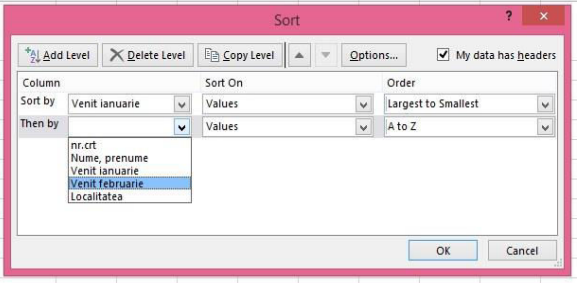
Sortarea tabelelor

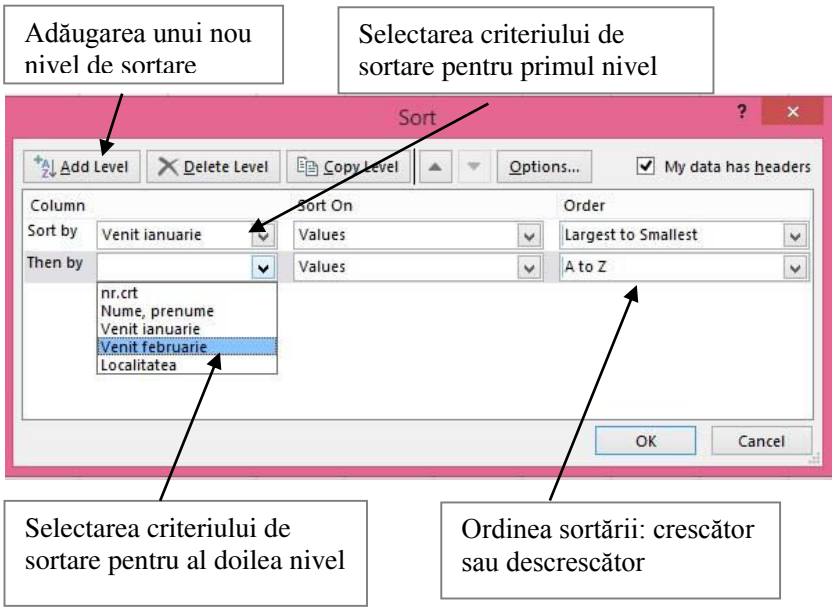
Pentru sortarea unui tabel trebuie să selectăm inițial datele pe care dorim să le sortăm. Dacă dorim să sortăm un tabel în funcție de informațiile dintr-o coloană, nu vom selecta doar coloana respectivă, ci întregul tabel, astfel datele vor rămâne corespunzătoare pe fiecare linie. Pentru a sorta datele accesăm **HOME>Editing>Sort&Filter.** Putem alege direct o sortare crescătoare (***Sort A to Z***) sau una descrescătoare (***Sort Z to A***), sau putem accesa ***Custom Sort*** care ne permite adăugarea unor criterii multiple.



De exemplu: Avem înregistrate veniturile în două luni consecutive și dorim să sortăm tabelul în funcție de venitul din luna ianuarie și apoi în funcție de cel din februarie. Vom accesa **HOME>*Cutom Sort***, unde vom adăuga ambele criterii de sortare.

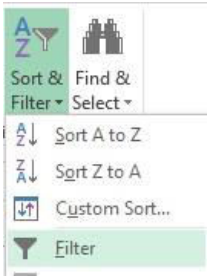
nr.crt	Nume, prenume	Venit ianuarie	Venit februarie	Localitatea
1	ARDELEAN MARINA	4546	3247	Brașov
2	BARNA VIOREL	4670	3900	Suceava
3	POP LUCIAN	3205	3546	Sibiu
4	BALAJ LAURA	2644	3626	Deva
5	FILIP ALIN	5392	4112	Oradea
6	ȘANDOR ANDRADA	2812	4663	București
7	BOT LORENA	1837	2956	Cluj-Napoca
8	COSMA MARIAN	2349	4454	Suceava
9	GABOR DANIEL	4265	5766	Timișoara
10	BOT FLORIN	1652	2347	București
11	BULZ AURORA	4058	2468	Alba-Iulia
12	IONESCU NATALIA	3545	4622	Baia Mare
13	LAZĂR EUGEN	3088	4469	București
14	BONACIU DORIN	3894	3969	Târgu-Mureș
15	MIHUȚ EDINA	3120	4997	Timișoara
16	OPREA ALEXANDRA	5479	5565	Suceava
17	AVRAM IONEL	3119	5650	Oradea





Filtrarea tabelelor

În Excel putem să "ascundem", să filtrăm, datele incluse într-un tabel în funcție tipul informațiilor pe care le conține: putem să ascundem un anumit tip de informații și să fie vizibil doar alt tip de informație. Pentru filtrarea unui tabel selectăm tabelul pe care dorim să aplicăm filtre și accesăm **HOME>Sort&Filter>Filter.** Înainte de a aplica filtrele, asigurați-vă că fiecare coloană are introdus un titlu (este realizat capul de tabel).



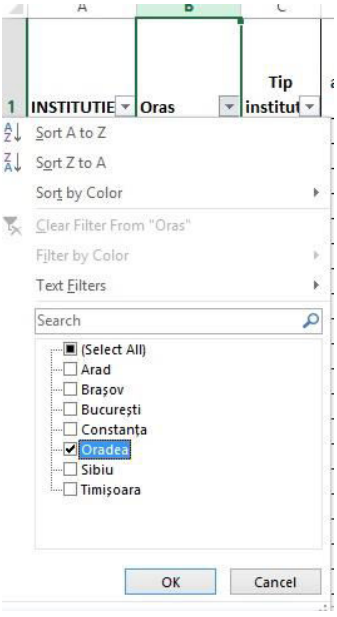
Vom observa că au fost adăugate butoane-săgeți pentru fiecare coloană a tabelului. Acestea apar în celulele din primul rând al tabelului (capul de tabel).

	A	B	C	D	E	F	G	H
1	INSTITUTIE	Oras	Tip institut	Nr angajati feme	Nr angajati barba	total angaja	Participare curs 1 formare profesiona	Participare curs 2 formare profesiona
2	Auto	Brașov	privat	6263	7852	**14115**	da	da
3	Artdecor	Oradea	privat	2498	9568	**12066**	nu	nu

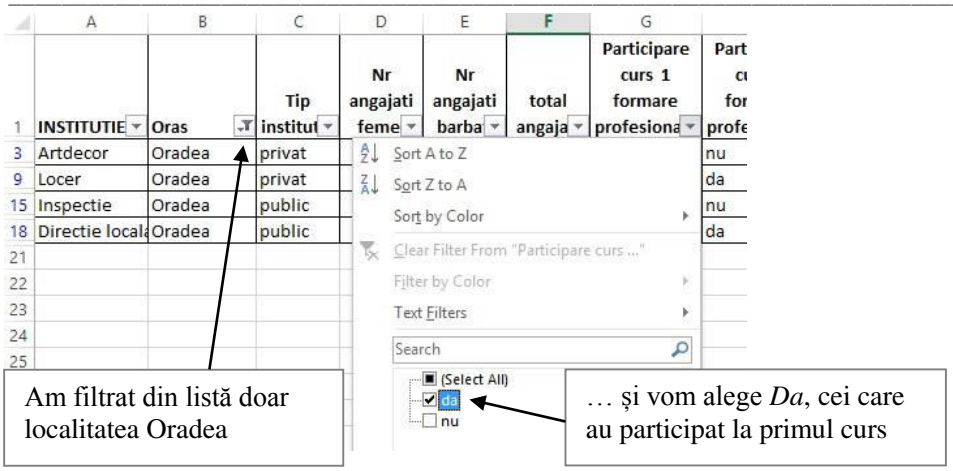
Accesând butonul-săgeată aferent fiecărei coloane putem să selectăm (“bifăm”) care dintre informațiile din listă să fie afișate ☑, respectiv care să fie ascunse ☐. După aplicarea unui filtru pentru o coloană, butonul nu va mai afișa săgeată ▼, ci o pâlnie care indică faptul că informația de pe coloana respectivă este filtrată.

De exemplu: Avem o bază de date cu angajații femei și bărbați ai unor organizații din diferite orașe (din domeniul privat sau public). Pentru fiecare organizație avem marcată situația dacă angajații au participat la două cursuri de formare.

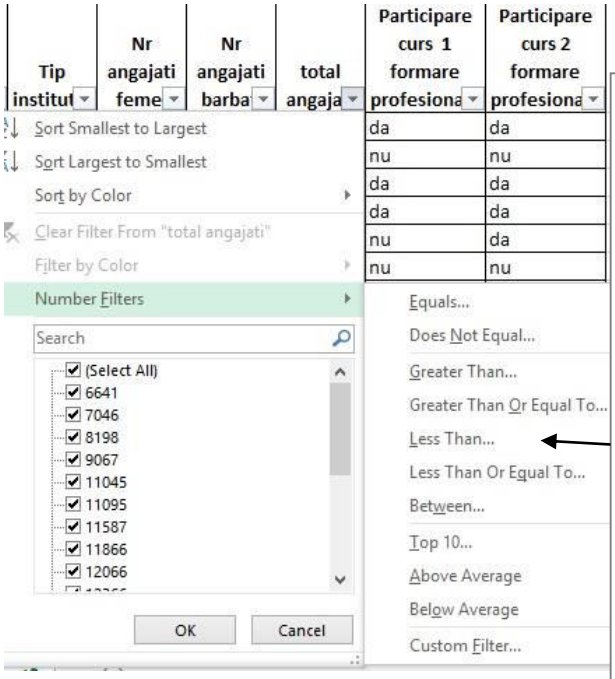
Dacă din întreaga bază de date de care dispunem dorim să vedem doar lista organizațiilor din Oradea, vom filtra în coloana care conține informațiile referitoare la localitate, localitatea Oradea.



Dacă de exemplu dorim să vedem lista organizațiilor din Oradea a căror angajați au participat la primul curs de formare, vom filtra pe de o parte în coloana care conține numele localităților, localitatea Oradea, iar în coloana referitoare la participarea la cursul 1 de formare vom selecta “da”.



În funcție de modul în care sunt introduse datele în tabel sau tipul datelor, putem să filtrăm celulele care nu conțin nicio informație (*Blanks*), sau să aplicăm filtre specifice datelor numerice:



Valori egale cu….

Valori care nu sunt egale cu…

Valori mai mari decât…

Valori mai mari sau egale decât…

Valori mai mici decât…

Valori mai mici sau egale decât…

Valori cuprinse într-un anumit interval

Primele 10 cele mai mari sau mai mici valori

Valori peste medie

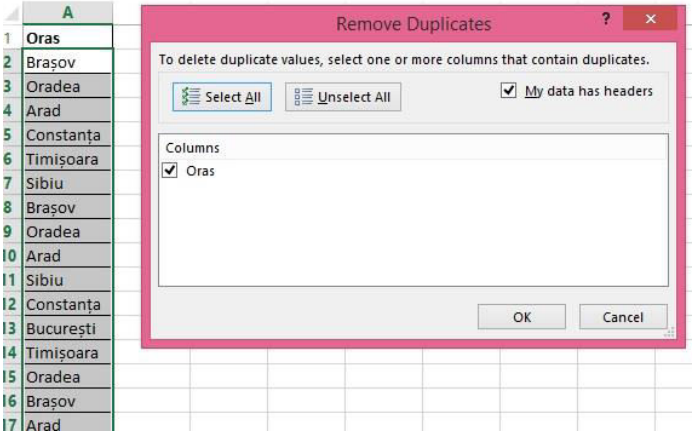
Valori sub medie

Pentru a vedea lista informațiilor introduse pe o coloană putem să tastăm ALT+ ↓

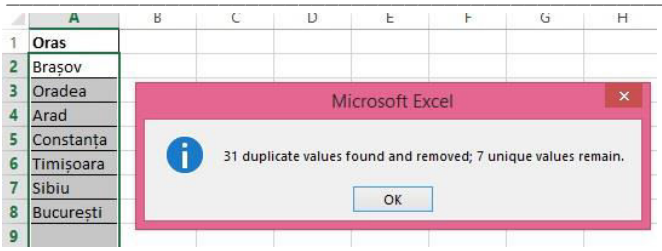
Eliminarea duplicatelor

Dacă lucrăm cu baze de date care conţin multe informații, de cele mai multe ori este necesară corectarea și gestiunea datelor. Sunt situații în care este necesar să eliminăm înregistrările multiple. Astfel, dacă o informație se multiplică, opțiunea ***Remove duplicates*** ne elimină informațiile din celule care se repetă și păstrează doar o singură intrare din aceeași informație.

De exemplu: Dacă avem o variabilă referitoare la localitate și dorim să aflăm care este lista localităților incluse în această variabilă, vom accesa **DATA>Remove duplicates**.



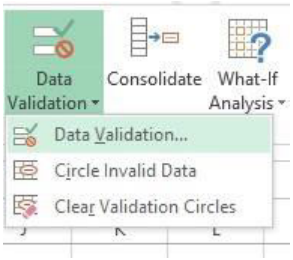
După ce vom confirma cu OK, în listă vor fi păstrate doar intrările singulare, iar Excel ne informează câte valori au fost șterse.

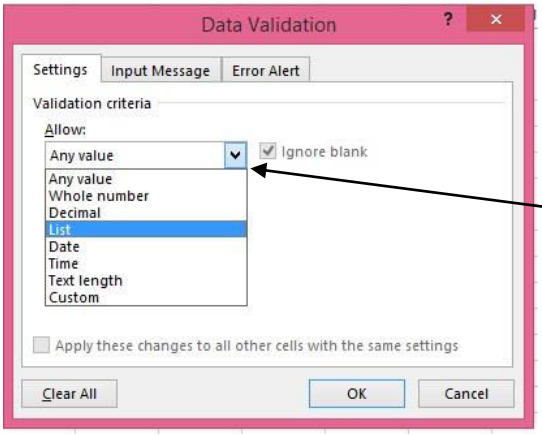


Trebuie să fim atenți când utilizăm această opțiune deoarece datele sunt șterse din baza de date. Dacă nu dorim ca acestea să fie șterse, fie vom opta pentru filtrarea datelor, fie vom folosi ***Conditional Formatting>Highlight Cell Rules>Duplicate Values***.

Validarea datelor

Când lucrăm cu baze de date realizate în Excel, putem să adăugăm restricții în ceea ce privește datele care pot fi introduse. În acest sens utilizăm anumite criterii pentru validarea datelor. Introducerea acestor criterii o realizăm accesând meniul **DATA>Data Validation**.





Sunt mai multe tipuri de criterii în funcție de care să fie validate datele:

număr întreg

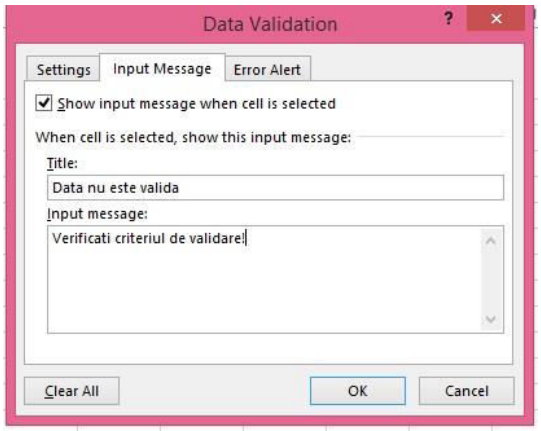
număr cu decimale

utilizarea unei liste cu itemi

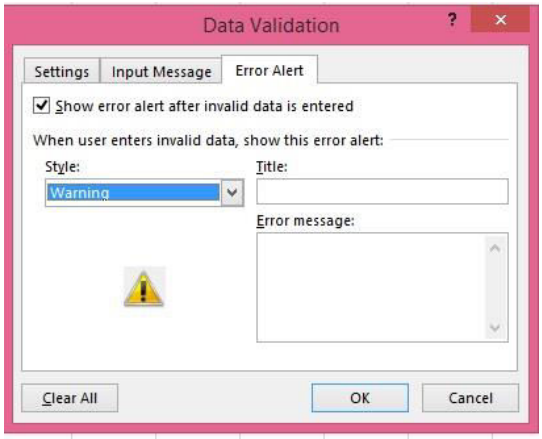
dată sau timp

lungimea textului

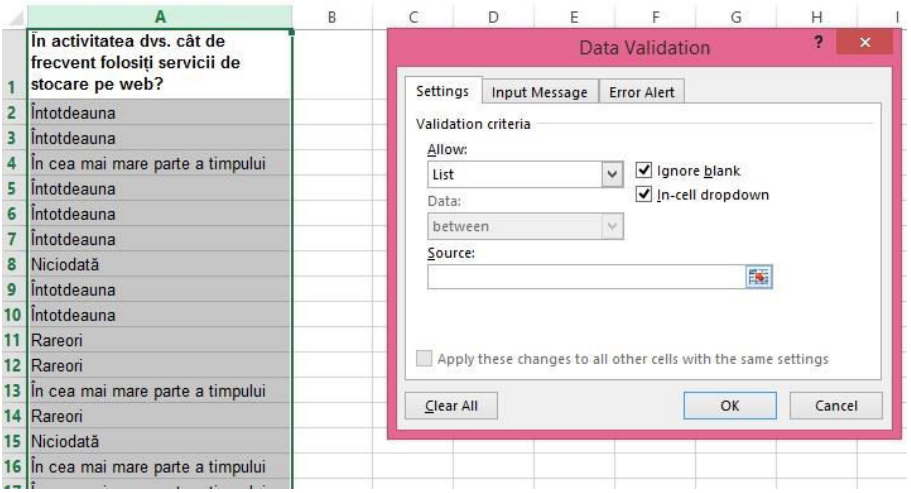
Putem opta pentru introducerea unui mesaj care va apărea în momentul în care datele introduse nu corespund criteriului de validare.



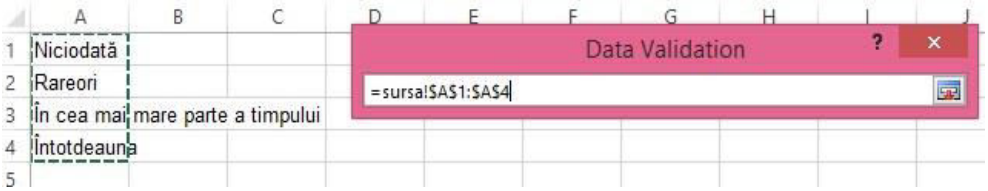
De asemenea, putem să adăugăm şi un avertisment de eroare:



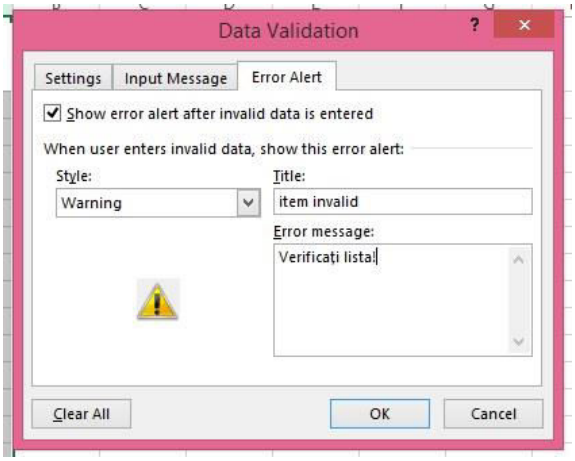
De exemplu: Dorim să introducem datele obținute într-un sondaj realizat în rândul angajaților din domeniul public. Avem următoarea întrebare *În activitatea dvs. cât de frecvent folosiți servicii de stocare pe web?*, cu variantele de răspuns *Niciodată, Rareori, În cea mai mare parte a timpului, Întotdeauna*. Când introducem răspunsurile primite de la subiecți, dorim să ne asigurăm că în baza de date vor fi adăugate doar cele 4 variante de răspuns şi nu vor apărea astfel alte erori. În acest sens vom adăuga un criteriu de validare a datelor introduse – lista celor 4 variante de răspuns.



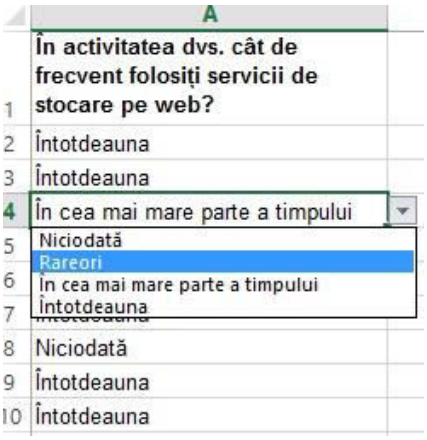
Pentru sursa listei (*Source*) vom selecta cele 4 variante de răspuns:



Vom adăuga și un mesaj-avertisment în situația în care sunt introduse date care nu corespund criteriului de validare:



Fiecare celulă va avea lista cu itemii disponibilă (accesând butonul), iar introducerea datelor poate fi realizată prin selecția unui item din listă.



Totodată, dacă introducem orice altă informație care nu se regăsește în listă, vom primi mesajul care ne avertizează că datele introduse nu sunt valide.

2 FORMULE ȘI FUNCȚII. ASPECTE INTRODUCTIVE

Excel este un program pentru calcul tabelar, ceea ce presupune că este folosit cel mai frecvent pentru aplicarea anumitor formule sau a unor funcții. Acestea utilizează datele introduse în foile de calcul.

Adăugarea formulelor și funcțiilor

Scrierea unei formule sau a unei funcții în Excel presupune urmarea unor pași foarte simpli:

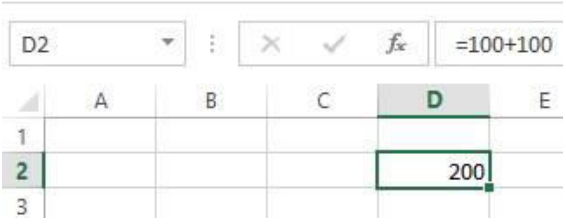
1. introducerea semnului egal (=) în celula în care dorim să realizăm un calcul

2. adăugarea expresiei (adăugarea tuturor elementelor specifice unei formule sau funcții)

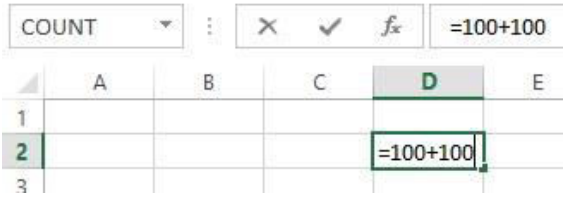
3. confirmarea cu ENTER

În momentul introduceri semnului egal =, programul intră în modul *Enter* (de introducere) a unei formule. Această situație presupune fie adăugarea manuală a textului (referințe sau operatori, folosind tastatura), fie selectarea referințelor (folosind cursorul, selectăm o celulă sau un grup de celule unde se află datele).

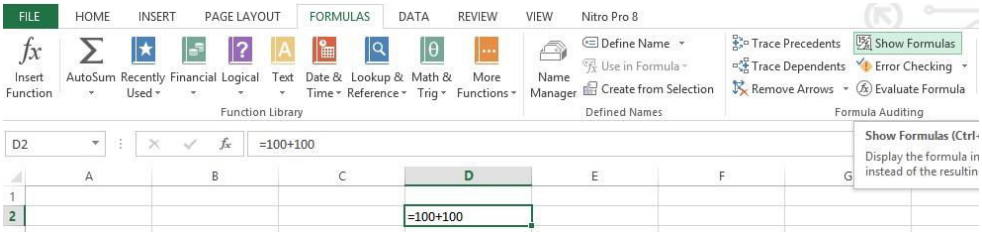
După confirmarea cu ENTER a formulei introduse, în celula în care a fost introdusă va apărea rezultatul. În bara destinată editării formulelor (marcată cu semnul f_x) va apărea formula care stă la baza calculului din celula respectivă.



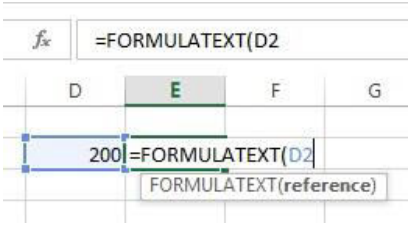
Pentru a edita o formulă introdusă într-o celulă, fie apăsăm dublu-click pe celula respectivă (pentru a intra în modul Editare), fie lucrăm direct în bara destinată editării formulelor, fie apăsăm F2 de pe tastatură (tot pentru a intra în modul Editare).

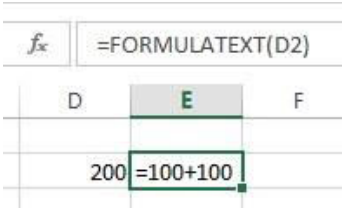


Putem să vizualizăm formulele introduse în celulele unei foi de calcul și accesând meniul **FORMULAS>Show Formulas.**



Tot pentru a vizualiza formulele introduse într-o celulă, putem să folosim funcția **=FORMULATEXT**(refernce). Această funcție ne returnează formula sub formă de text.





Elementele formulelor și funcțiilor

O formulă include referințele (celula sau celulele în care se află valorile incluse în formulă), operatorii aritmetici (adunare, scădere etc.) și/sau constante (o valoare sau mai multe, adăugate direct în formulă). O formulă poate să conțină și funcții, iar o funcție poate să fie construită din mai multe funcții (unul sau mai multe argumente ale unei funcții să fie reprezentate de către alte funcții).

Introducerea unei funcții presupune după adăugarea semnului "=" scrierea numelui funcției respective, urmată obligatoriu de paranteze în care sunt introduse argumentele funcției **=FUNCȚIA**(…). O funcție poate să conțină unul sau mai multe argumente. În situația în care sunt adăugate mai multe argumente, acestea sunt separate între ele de semnul virgulă (,) sau punct-și-virgulă (;). Argumentele unei funcții pot să fie obligatorii sau opționale. Dacă argumentele introduse nu sunt numerice sau nu sunt selectate direct din referințele din foaia de calcul, cu alte

cuvinte dacă lucrăm cu un argument sub formă de text sau simbol, acestea trebuie introduse în ghilimele ("…").

Exemple de formule și funcții	
=A1+A2	Adună valoarea din celula A1 și valoarea din celula A2 *Referințe*: A1 și A2 *Operator*: +
=(C1*C4)/2	Returnează valoarea înmulțirii valorilor din celulele C1 și C4, împărțit la 2 *Referințe*: C1 și C4 *Operatori*: * și / *Constanta*: 2
=SUM(B1:B5)^2	Returnează suma valorilor incluse în grupul de celule de pe coloana B, de la rândul 1 la rândul 5, valoare ridicată la pătrat *Referințe*: grup de celule B1:B5 *Operator*: ^ *Funcție*: SUM()
=SUM(MAX(D2:D7), MAX(E2:E7))*3	Returnează suma dintre valoarea maximă regăsită în grupul de celule D2:D7 și valoarea maximă din grupul de celule E2:E7, rezultatul sumei înmulțit cu 3 *Referințe*: grup de celule D2:D7 și E2:E7 *Operator*:* *Funcții*: SUM() și MAX()

Operatori matematici utilizați în cadrul formulelor

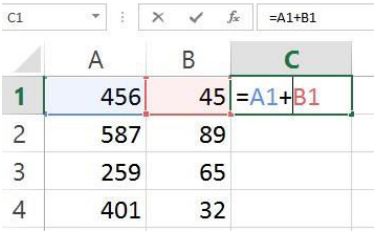
Operator	**Descriere**	**Exemplu**
+	Adunare	=1+1 (rezultat 2)
-	Scădere	=3-1 (rezultat 2)
*	Înmulțire	=5*5 (rezultat 25)
/	Împărțire	=10/2 (rezultat 5)
^	Exponențial (ridicare la puterea specificată)	=3^2 (rezultat 9)
%	Procent	=8% (rezultat 0.08)

Atenție: țineți cont de ordinea operațiilor matematice. Adăugați paranteze pentru a specifica ordinea operațiilor.

Opțiunea auto-umplere

Atunci când introducem formule în celule este important să utilizăm referințe. Cu alte cuvinte, nu vom introduce direct valorile pe care dorim să le includem în formulă, ci vom face trimiteri la celulele care conțin valorile respective.

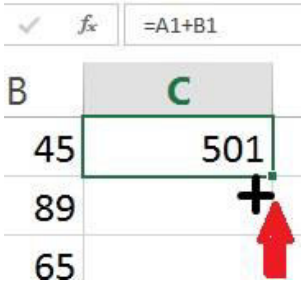
De exemplu: Dacă dorim să adunăm valorile introduse în coloana A și coloana B, nu vom scrie valorile direct în formulă "=456+45", ci vom utiliza referințele (trimiteri directe la valorile incluse în celula A1, respectiv B1).



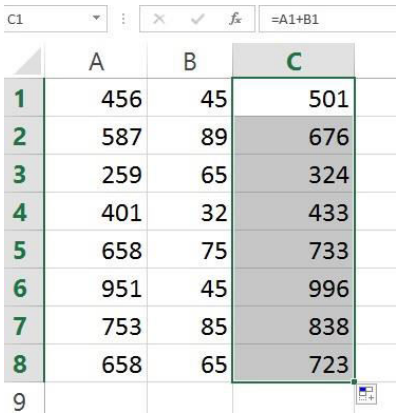
Acest lucru ne permite pe de o parte să corectăm direct rezultatele obținute în urma aplicării unei formule (sau funcții) în situația în care modificăm valorile conținute de formula (sau funcția) respectivă, iar pe de altă parte, putem să utilizăm opțiunea auto-umplere care asigură multiplicarea formulei (sau funcției) și astfel, scurtează considerabil timpul de calcul.

Pentru utilizarea opțiunii auto-umplere, vom selecta celula în care este introdusă formula (funcția). Dacă poziționăm cursorul (*mouse*-ul) pe pătrățelul din marginea celulei - colțul din dreapta-jos, acesta se va transforma în semnul plus "**+**". Pentru a multiplica formula, fie tragem cursorul până selectăm toate celulele unde dorim să multiplicăm formula, fie apăsăm dublu-click când cursorul este sub forma "**+**". Această ultimă variantă, multiplică formula în dreptul tuturor liniilor în care sunt introduse valori.

De exemplu: Dacă în coloana C dorim să calculăm suma valorilor incluse în coloana A și coloana B, este suficient să introducem o singură dată formula în celula C1, apoi pentru restul coloanei vom folosi opțiunea auto-umplere.



Am apăsat dublu-click pe colțul din dreapta-jos al celulei C1 (când cursorul este în formă de +), iar coloana C este auto-umplută cu rezultatele obținute în urma adunării valorilor din coloana A și coloana B.

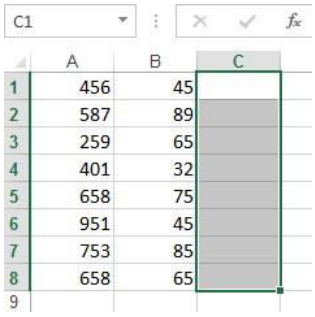


	A	B	C
1	456	45	501
2	587	89	676
3	259	65	324
4	401	32	433
5	658	75	733
6	951	45	996
7	753	85	838
8	658	65	723
9			

Pentru auto-umplerea formulelor în celule putem să folosim și combinația de taste **CTRL+Enter**.

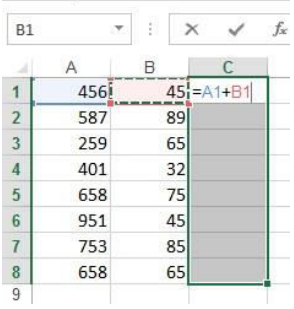
De exemplu: Pentru datele prezentate anterior, înainte de a introduce formula în celula C1, vom selecta toate celulele din coloana C unde dorim să calculăm sumele. Apoi vom scrie formula =A1+B1 și vom apăsa tastele CRTL și Enter, iar astfel în toate celulele selectate anterior vor apărea rezultatele formulei.

Pasul 1: Selectăm



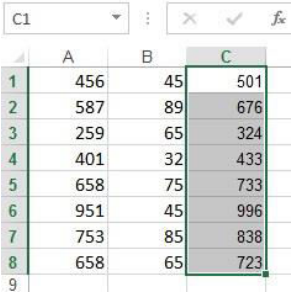
	A	B	C
1	456	45	
2	587	89	
3	259	65	
4	401	32	
5	658	75	
6	951	45	
7	753	85	
8	658	65	
9			

Pasul 2: Scriem formula



	A	B	C
1	456	45	=A1+B1
2	587	89	
3	259	65	
4	401	32	
5	658	75	
6	951	45	
7	753	85	
8	658	65	
9			

Pasul 3. CTRL+Enter



	A	B	C
1	456	45	501
2	587	89	676
3	259	65	324
4	401	32	433
5	658	75	733
6	951	45	996
7	753	85	838
8	658	65	723
9			

Pentru a multiplica o formulă sau o funcție putem să apăsăm dublu-click când cursorul este sub forma semnului "**+**". Acest mod de lucru va auto-umple toate celulele în dreptul cărora se află valori pentru care se poate aplica formula (sau funcția).

Utilizarea referințelor

O formulă (sau o funcție) va include întotdeauna cel puțin o referință. Referințele sunt acele celule sau grup de celule care indică poziția în foaia de calcul (*sheet*) a datelor ce sunt incluse într-o formulă (sau funcție). Referințele se pot afla în aceeași foaie de calcul, într-o altă foaie de calcul din același fișier sau chiar în foi de calcul din alte fișiere față de cel în care se află introdusă formula (funcția) respectivă.

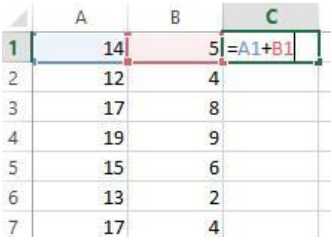
Referințele pot să facă trimitere la informația cuprinsă într-o singură celulă sau la datele incluse în mai multe celule. Atunci când lucrăm cu un grup de celule, semnul utilizat este ":".

Exemple: selecție o celulă sau un grup de celule
A3
Referința este informația inclusă în celula A3
A1:A20
Selecția informațiilor incluse în celulele din coloana A, de la rândul 1 la rândul 20
A1:F1
Selecția informațiilor incluse în celulele de pe rândul 1, de la coloana A la coloana F
C1:G35
Selecția grupului de celule de la rândul 1 la 35 și de la coloana C la coloana G

O formulă sau o funcție poate să conțină referințe relative sau referințe absolute.

Referințele relative se referă la poziția relativă a celulei la care se face referire în celula care conține formula sau funcția. Cu alte cuvinte, poziția celulei care este referință într-o formulă sau funcție se poate modifica. Dacă multiplicăm formula (sau funcția) pe coloane sau rânduri (folosind opțiunea de auto-umplere), celulele de referință se vor modifica corespunzător acestei multiplicări.

De exemplu: Dacă în coloana C dorim să adunăm datele din celulele A1 și B1, vom folosi în formula introdusă în celula C1 referințele relative A1 și B1. În acest fel, ne vom asigura că după ce vom multiplica formula pe coloana C, referințele din celulele A1 și B1 se vor modifica corespunzător fiecărui rând.



Pozițiile celulelor A1, respectiv B1 sunt relative. Acestea se vor modifica dacă vom multiplica formula în jos pe coloana C (auto-umplere). A1 se va modifica în A2, A3 etc., iar B1 în B2, B3 etc. (corespunzător formulei multiplicate pe rândul 2, rândul 3 etc.).

	A	B	C
1	14	5	19
2	12	4	16
3	17	8	=A3+B3
4	19	9	28
5	15	6	21

Este important ca atunci când utilizați referințe relative într-o formulă și faceți modificări la un moment dat în conținutul formulei, trebuie să reluați multiplicarea formulei pe coloană sau pe rând, ca această modificare să se aplice tuturor celulelor care conțin referințele.

Referințele absolute nu permit ajustarea poziției unei celule inclusă într-o formulă (sau funcție). Utilizarea semnului $ asigură blocarea poziției celulei respective. Deoarece poziția unei celule ține cont de intersecția dintre o coloană și un rând, blocarea acesteia poate fi realizată pe rând, pe coloană (referințe mixte) sau pe rând și coloană (referințe absolute). Adăugarea semnului $ în fața literei care marchează coloana blochează poziția pe coloană. Adăugarea semnului $ în fața cifrei care marchează rândul blochează poziția pe rând. Iar adăugarea semnului $ în fața literei coloanei și în fața cifrei rândului blochează poziția pe coloană și rând.

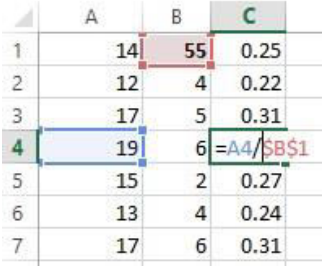
De exemplu:

A$5 – este blocată poziția pe rândul 5, nu dar nu este blocată poziția pe coloană.

$C6 – este blocată poziția pe coloana C, nu dar nu este blocată poziția pe rând.

$B$5 – este blocată poziția pe coloana B, rândul 5.

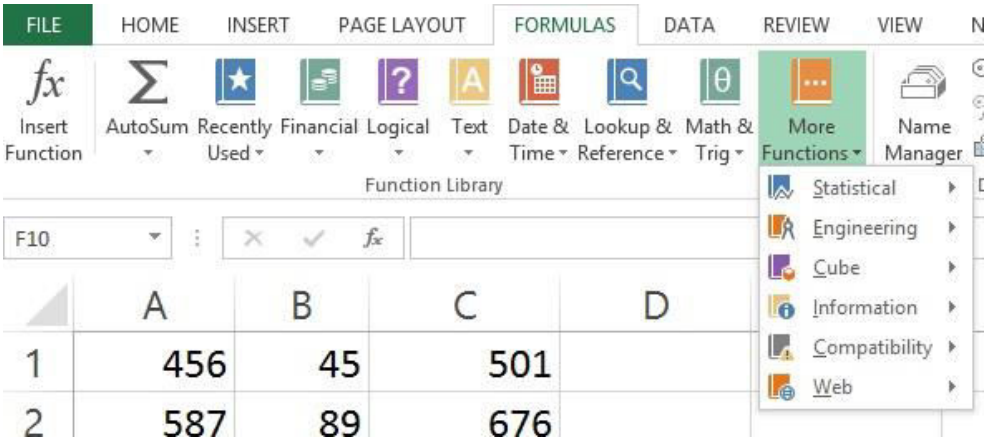
De exemplu: Dacă dorim în coloana C să obținem rezultatele raportării datelor din coloana A la valoarea inclusă în celula B1, vom bloca poziția celulei B1. Astfel, dacă vom multiplica formula pe coloana C în jos, se va modifica poziția relativă A1, dar poziția B1 va rămâne întotdeauna aceeași.



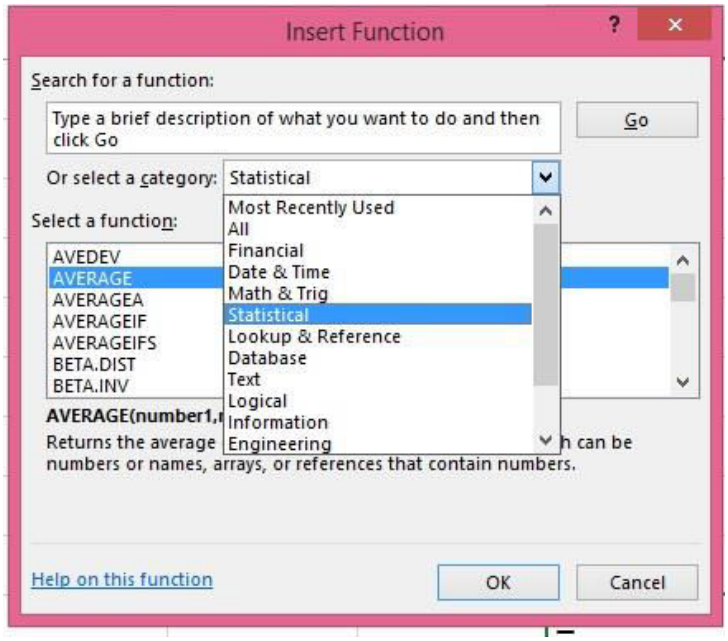
	A	B	C
1	14	55	0.25
2	12	4	0.22
3	17	5	0.31
4	19	6	=A4/$B$1
5	15	2	0.27
6	13	4	0.24
7	17	6	0.31

Introducerea funcțiilor

Funcțiile sunt formule predefinite care permit efectuarea calculelor folosind argumente într-o ordine prestabilită. Funcțiile pot fi accesate din meniul **FORMULAS**. Aici sunt disponibile principalele categorii de funcții: Funcții Financiare, Funcțiile Logice, Funcții Text, Funcții pentru Date și Timp, Funcții de Căutare și de Referință, Funcții Matematice și Trigonometrice, Funcțiile Statistice, Funcții de Inginerie, Funcții de Cub, Funcții de Informații, Funcții de Compatibilitate, Funcțiile Web.

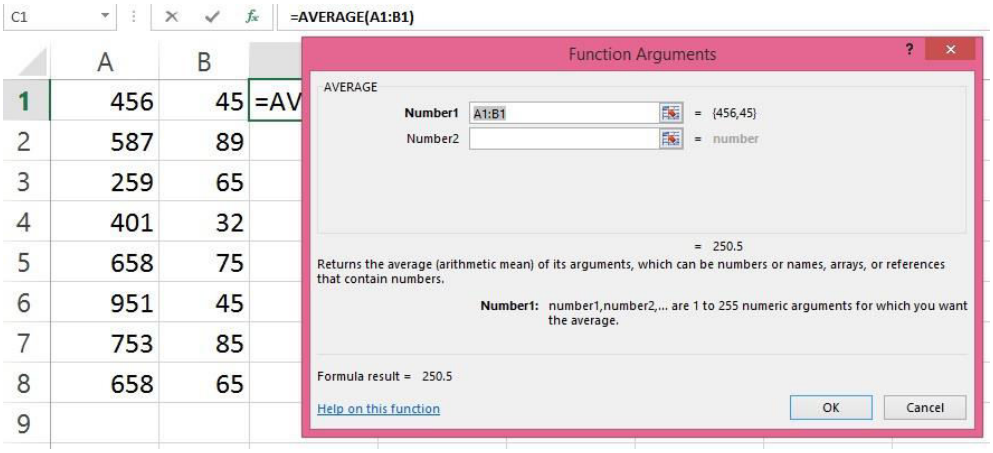


Putem să introducem o funcție fie accesând meniul **FORMULAS>Insert Function**, fie prin introducerea semnului "=" în celula în care dorim să adăugăm funcția și scrierea numelui funcției. În prima situație, prin accesarea meniului **FORMULAS>Insert Function** se deschide fereastra în care sunt afișate categoriile funcțiilor și lista tuturor funcțiilor din categoriile respective. Dacă vom selecta o funcție din listă, fereastra va include și o descriere a funcției respective.



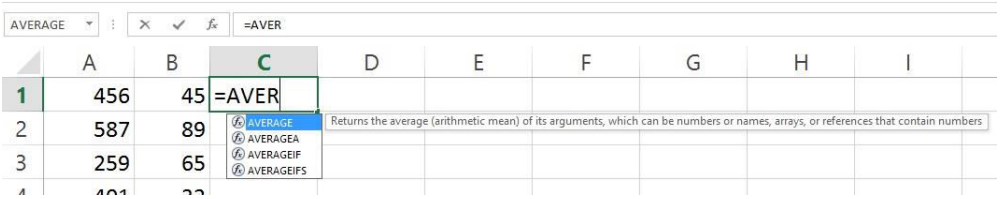
De exemplu: Dacă dorim să calculăm media aritmetică dintre valorile incluse în coloana A și coloana B, vom selecta funcția **=AVERAGE**. Pe parcursul acestei cărți, sunt prezentate principalele funcții utilizate în prelucrarea și analiza datelor.

După ce am confirmat cu OK funcția selectată, în fereastra următoare trebuie să adăugăm argumentele funcției. În cazul exemplului nostru, vom adăuga referința A1:B1. Apoi vom confirma comanda cu OK.

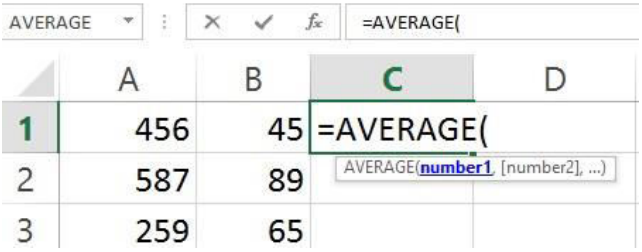


Unul dintre principalele avantaje în utilizarea programului Excel constă în recunoașterea automată a funcțiilor. Este suficient să adăugăm după semnul "=" primele litere din funcția căutată, iar Excel ne afișează lista tuturor funcțiilor a căror

nume începe cu literele respective. Mai mult, orice funcție selectată va fi însoțită de un text care descrie funcția respectivă.

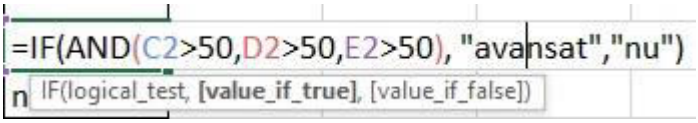


După introducerea numelui funcției, vom adăuga întotdeauna paranteză deschisă "(", iar în acest fel trecem la adăugarea argumentelor și a referințelor pentru funcția respectivă.

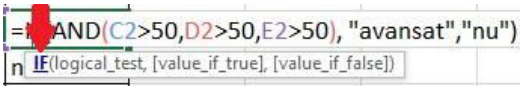


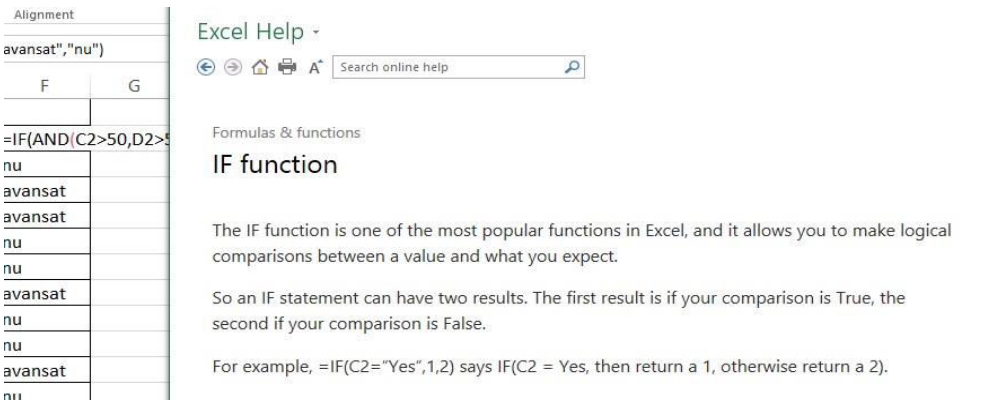
Atunci când introducem argumentele unei funcții, Excel afișează întotdeauna o casetă-text (*ScreenTip*) care este extrem de folositoare deoarece ne indică poziția în care ne aflăm cu editarea formulei. Argumentul pe care îl editățm în acel moment este marcat cu *Bold*. Această casetă-text poate fi mutată într-o altă poziție dacă se suprapune pe o celulă de care avem nevoie (mutăm folosind mouse-ul când cursorul este).

De exemplu: Caseta- text (*ScreenTip*) de mai jos ne indică faptul că referința introdusă în acel moment ("avansat") se referă la [*value_if_true*], adică valoarea care va fi folosită în cazul în care condiția testată prin funcția IF (DACĂ) este adevărată.

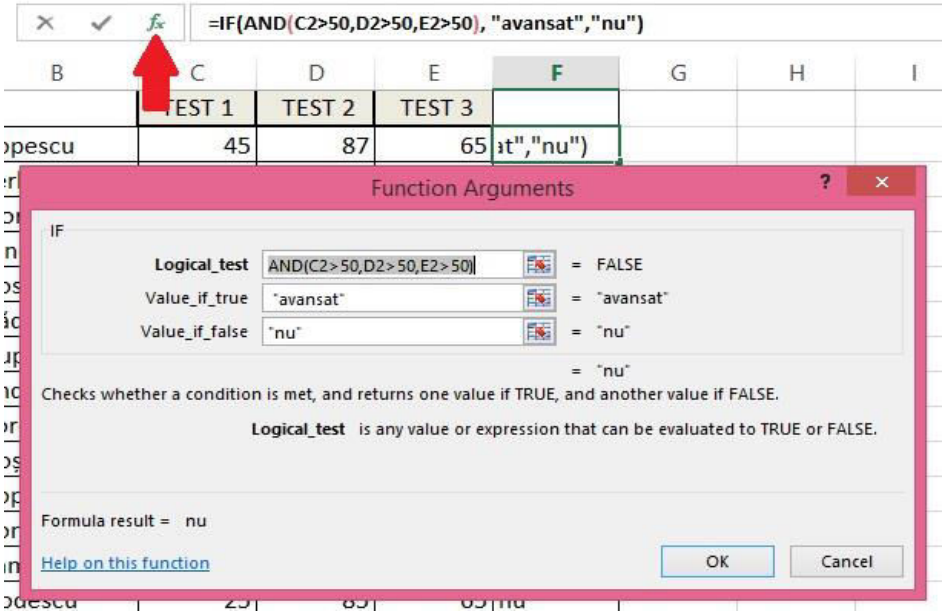


În situația în care nu cunoaștem suficient de bine funcția selectată, putem să apăsăm cu mouse-ul pe numele funcției din caseta-text. În acest fel se va deschide meniul **HELP** care cuprinde o descriere a funcției respective.




Alignment
avansat","nu")
F
G
=IF(AND(C2>50,D2>5
nu
avansat
avansat
nu
nu
avansat
nu
nu
avansat
Excel Help
Search online help
Formulas & functions
IF function
The IF function is one of the most popular functions in Excel, and it allows you to make logical comparisons between a value and what you expect.
So an IF statement can have two results. The first result is if your comparison is True, the second if your comparison is False.
For example, =IF(C2="Yes",1,2) says IF(C2 = Yes, then return a 1, otherwise return a 2).

Putem să verificăm detaliile privind argumentele conținute de o formulă accesând direct funcția dând click pe semnul f_x din bara de sus. Astfel deschidem fereastra care cuprinde toate detaliile aferente formulei: o descriere a funcției, argumentele adăugate, rezultatul obținut.



ATENȚIE: Excel va calcula în locul nostru, dar comenzile noi trebuie să le scriem corect. Excel nu va "gândi" în locul nostru care sunt formulele sau care sunt funcțiile potrivite pentru obținerea unor rezultate.

Detaliile privind principalele funcții utilizate în prelucrarea și analiza datelor sunt prezentate în următoarele secțiuni ale cărții.

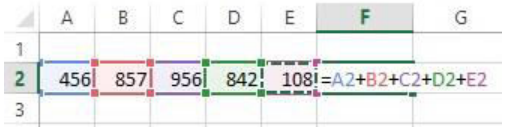
3 FUNCȚIILE MATEMATICE ȘI STATISTICE

SUM, PRODUCT, SUMPRODUCT, AVERAGE, COUNT, COUNTA, COUNTBLANCK, MIN, MAX, SUBTOTAL

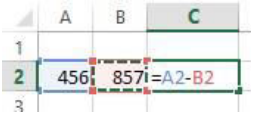
Formulele aritmetice simple sunt probabil cele mai utilizate operațiuni în Excel. Acestea presupun adunări, scăderi, înmulțiri și împărțiri. De regulă pentru aceste tipuri de calcule sunt folosite direct referințele celulelor și operatorii corespunzători: "**+**", "**-**", "*****", "**/**".

De exemplu:

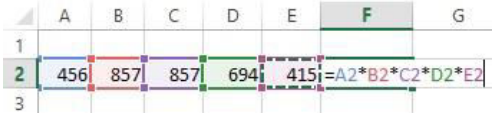
Adunare:



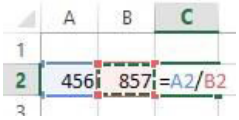
Scădere:



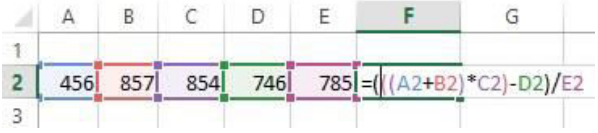
Înmulțire:



Împărțire:



Toate operațiile într-o singură formulă:



Dar ce facem atunci când dorim să însumăm sau să înmulțim o serie mai lungă de date? Adăugarea referințelor și operatorilor ar consuma foarte mult timp. În astfel de situații este de preferat să folosim funcțiile. Funcțiile pot fi introduse fie accesând meniul **FORMULAS>Insert functions**, fie introducând direct numele funcției în

celula în care dorim să realizăm calculul. De regulă, vom folosi cea de-a două variantă deoarece este mai rapidă.

Funcția
=SUM(number1, [number2], …)

Descriere

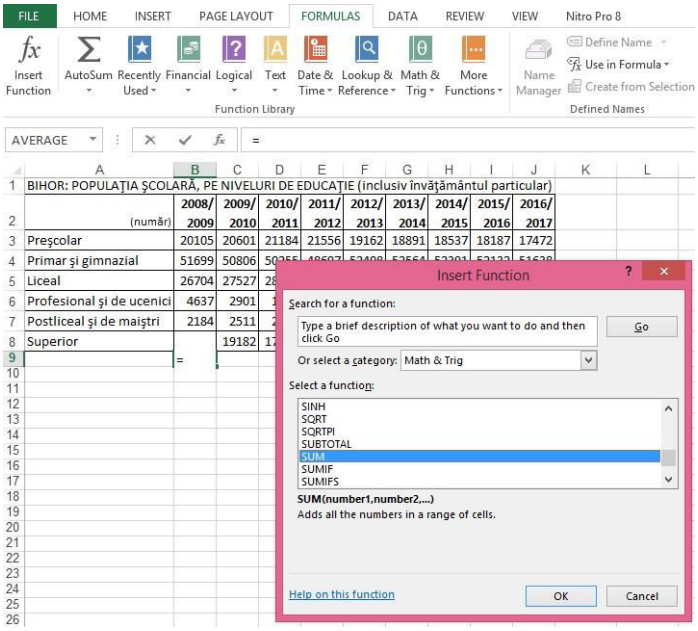
Funcția =SUM însumează o serie de date. Funcția se regăsește în categoria funcțiilor Matematice și trigonometrice.

Argumentele

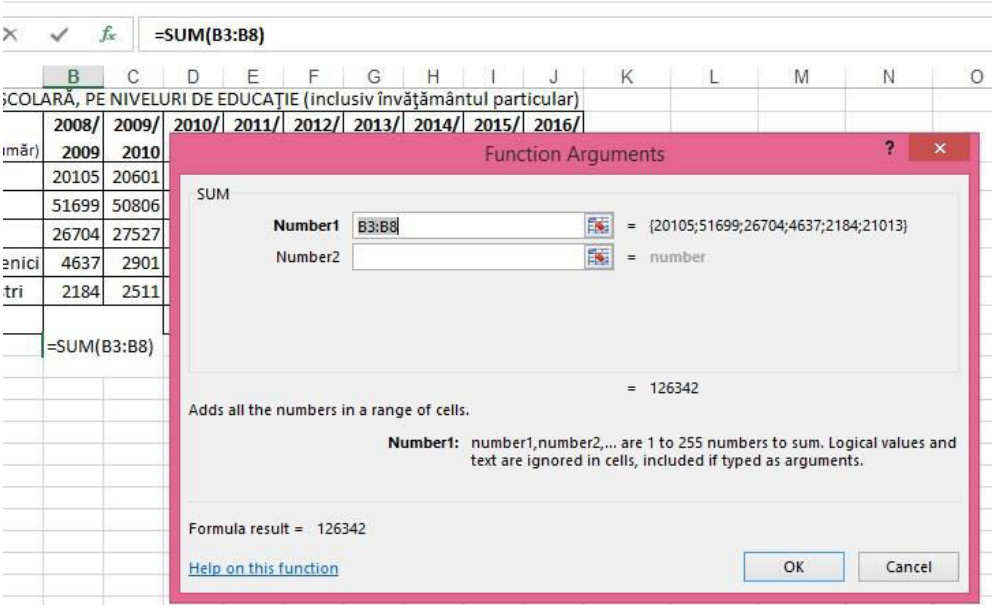
Number1…[numer2] – valorile sau referințele care conțin valorile pe care dorim să le însumăm.

De exemplu: Avem distribuția numărului de persoane înregistrate pe niveluri de învățământ (sursa datelor: Direcţia Judeţeană de Statistică BIHOR, www.bihor.insse.ro). Vom folosi funcția **=SUM** pentru a calcula numărul total de persoane înscrise pe fiecare an școlar. Sunt prezentate cele patru modalităţi de adăugare a funcției **=SUM**, cu mențiunea că este recomandată fie introducerea numelui funcției direct în celula în care dorim să obținem rezultatul, fie folosirea combinației de taste.

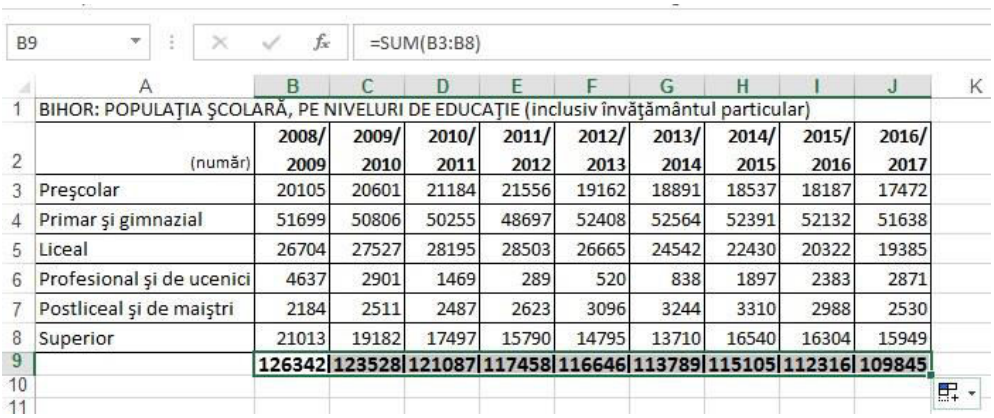
Selectăm celula în care dorim să obținem rezultatul și accesăm meniul **FORMULAS>Insert functions.** Selectăm categoria *Math & Trig*, iar din listă vom selecta funcția **SUM**.



În fereastra următoare vom introduce argumentele, și anume selecția celulelor în care sunt introduse datele pe care dorim să le însumăm (în exemplul nostru B3:B8). Confirmăm cu OK, iar rezultatul va apărea în celula în care am inserat funcția.


=SUM(B3:B8)
ȘCOLARĂ, PE NIVELURI DE EDUCAȚIE (inclusiv învățământul particular)
Function Arguments
SUM
Number1 B3:B8 = {20105;51699;26704;4637;2184;21013}
Number2 = number
= 126342
Adds all the numbers in a range of cells.
Number1: number1,number2,... are 1 to 255 numbers to sum. Logical values and text are ignored in cells, included if typed as arguments.
Formula result = 126342
Help on this function
OK
Cancel

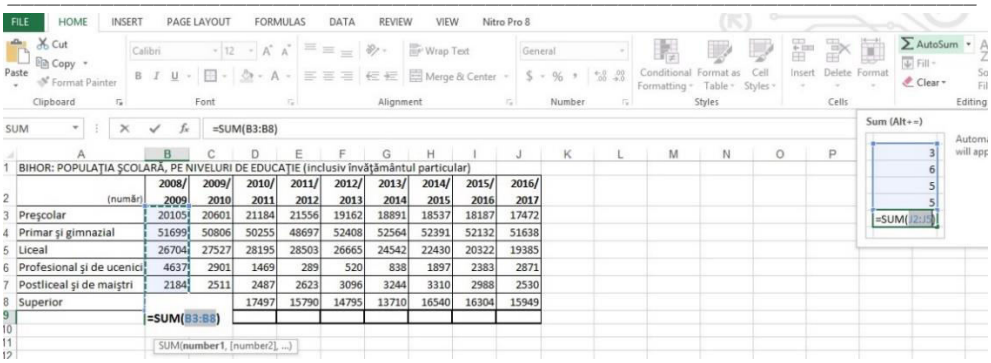
Vom folosi opțiunea auto-umplere și vom multiplica formula pentru toate coloanele, și anume pentru toți anii școlari.



	A	B	C	D	E	F	G	H	I	J
1	BIHOR: POPULAȚIA ȘCOLARĂ, PE NIVELURI DE EDUCAȚIE (inclusiv învățământul particular)									
2	(număr)	2008/ 2009	2009/ 2010	2010/ 2011	2011/ 2012	2012/ 2013	2013/ 2014	2014/ 2015	2015/ 2016	2016/ 2017
3	Preşcolar	20105	20601	21184	21556	19162	18891	18537	18187	17472
4	Primar şi gimnazial	51699	50806	50255	48697	52408	52564	52391	52132	51638
5	Liceal	26704	27527	28195	28503	26665	24542	22430	20322	19385
6	Profesional şi de ucenici	4637	2901	1469	289	520	838	1897	2383	2871
7	Postliceal şi de maiştri	2184	2511	2487	2623	3096	3244	3310	2988	2530
8	Superior	21013	19182	17497	15790	14795	13710	16540	16304	15949
9		126342	123528	121087	117458	116646	113789	115105	112316	109845
10										
11										

Vom citi rezultatul astfel: în sistemul de învățământ, în anul școlar 2008/2009 erau înscriși în total 126342 de elevi și studenți.

O altă modalitate prin care putem să calculăm sumele este accesarea din meniul **HOME** a butonului **AutoSum** care are simbolul pentru sumă ∑. Pașii sunt următorii: click pe semnul ∑, verificăm dacă selecția celulelor este corectă, click pe semnul ∑ pentru confirmare.



A treia modalitate prin care putem să introducem funcția **=SUM**, este scrierea directă a funcției în celula în care dorim să obținem rezultatul. Putem să scriem funcția și în bara destinată funcțiilor marcată cu semnul f_x.

Pentru exemplul nostru vom scrie: **=SUM**(B3:B8). Pentru adăugarea argumentelor, putem fie să le scriem, fie să selectăm direct (folosind mouse-ul) celulele în care se află valorile pe care dorim să le însumăm.

Probabil cel mai rapid mod în care putem să adăugăm funcția **=SUM** este folosirea combinației de taste **Alt=**. Apăsând cele două taste deodată este introdusă automat funcția =SUM, la care trebuie adăugate argumentele (pentru exemplul nostru B3:B8), fie prin scrierea lor, fie prin selectarea lor.

Funcția

=PRODUCT(number1, [number2], …)

Descriere

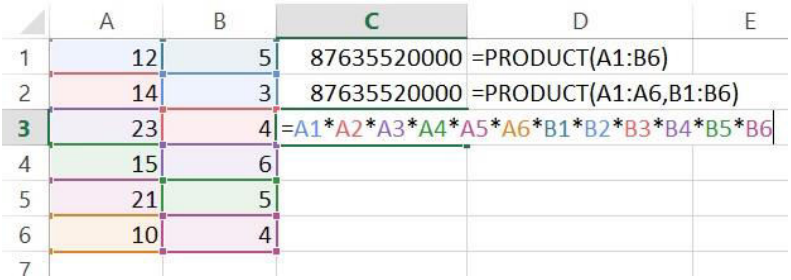
Funcția **=PRODUCT** (tot din categoria funcțiilor Matematice și trigonometrice) înmulțește toate valorile incluse într-o serie de date.

Argumentele

Number1 …[numer2] – valorile sau referințele care conțin valorile pe care dorim să le înmulțim.

Putem să inserăm funcția accesând meniul **FORMULAS>Insert functions,** fie să scriem direct numele funcției în celula în care dorim să obținem rezultatul. Funcția este utilă când avem multe cifre pe care trebuie să le înmulțim, iar utilizarea operatorului "*" ar presupune un consum mare de timp.

De exemplu: Este mai simplu să scriem direct funcția =**PRODUCT**(A1:B6) decât să adăugăm fiecare valoare din celule și să folosim operatorul "*". În exemplul de mai jos sunt prezentate trei formule care calculează același lucru, și anume înmulțirea tuturor cifrelor din coloanele A și B.



Funcția

=SUMPRODUCT(array1, [array2], [array3] …)

Descriere

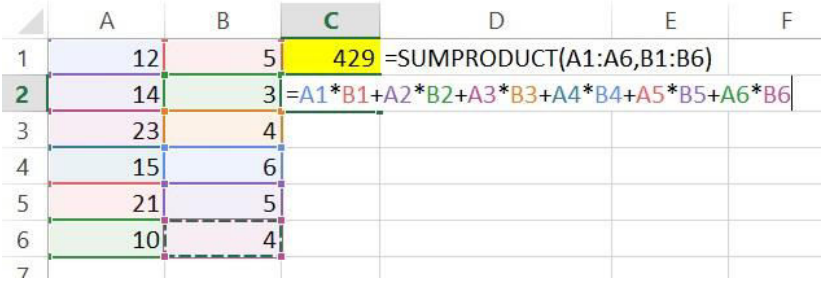
Funcția **=SUMPRODUCT** (din categoria funcțiilor Matematice și trigonometrice) returnează suma valorilor înmulțite dintr-o matrice de date.

Argumentele

Array1– seria de date pe care dorim să le înmulțim și să le adunăm apoi.

Array1,2…– următoarele serii de date pe care dorim să le înmulți și să le adunăm.

De exemplu: Necesită mai puțin timp să scriem funcția **=SUMPRODUCT**(A1:A6,B1:B6), decât să utilizăm operatorii "+" și "*" dacă dorim să calculăm suma valorilor obținute prin înmulțirea numerelor din coloana A cu cele din coloana B.



Totodată, putem să obținem rezultatul de mai sus, dacă vom realiza calculele în două etape: prima dată vom calcula înmulțirile, apoi vom însuma rezultatele. Totuși, cu funcția **=SUMPRODUCT** facem cea mai mare economie de timp.

	A	B	C
1	12	5	**60**
2	14	3	**42**
3	23	4	**92**
4	15	6	**90**
5	21	5	**105**
6	10	4	**40**
7			**=SUM(C1:C6)**

Funcția

=AVERAGE(number1, [number2], …)

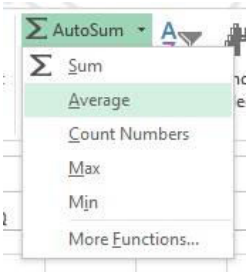
Descriere

Cu funcția **=AVERAGE** calculăm mediile aritmetice. Funcția se regăsește în categoria funcțiilor statistice. Detalii privind analiza statistică a datelor și modul de calcul a principalilor indicatori statistici se regăsesc într-o altă secțiune a cărții. Totuși, calcularea mediilor, alături de sume, sunt principalele operațiuni realizate folosind Excel.

Argumentele

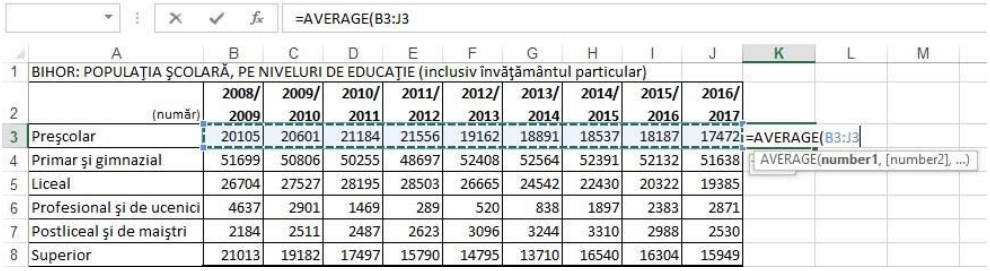
Number1…[numer2] – valorile sau referințele care conțin valorile pe baza cărora dorim să calculăm media aritmetică.

De exemplu: Pentru datele prezentate anterior privind numărul de persoane înscrise pe niveluri de învățământ, vom calcula în medie câte persoane au fost înscrise în perioada 2008-2016. Vom folosi funcția **=AVERAGE**. Putem să optăm pentru oricare dintre variantele privind introducerea funcției prezentate mai sus. Astfel, fie introducem funcția accesând meniul **FORMULAS>Insert functions**, fie accesăm din meniul **HOME**, funcția AVERAGE (în listă este următoarea după funcția SUM), fie scriem direct funcția în celula în care dorim să obținem rezultatul.



După ce am adăugat direct numele funcției în celula în care dorim să obținem rezultatul, adăugăm (prin selecție) referințele (în exemplul nostru B3:J3), apoi

confirmăm apăsând tasta ENTER. Vom folosi din nou opțiunea de auto-umplere pentru multiplicarea funcției pentru toate rândurile.



Funcția

=COUNT(value1, [value2], …)

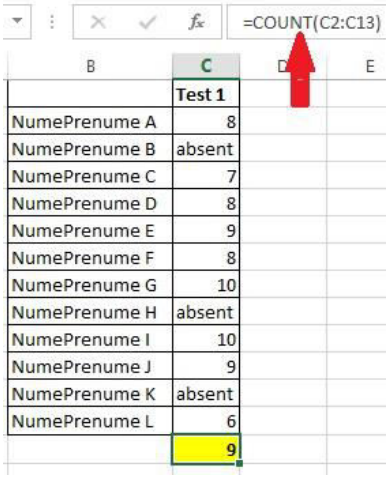
Descriere

Funcția **=COUNT** (din categoria funcțiilor Statistice) returnează numărul de celule dintr-un șir care conțin numere.

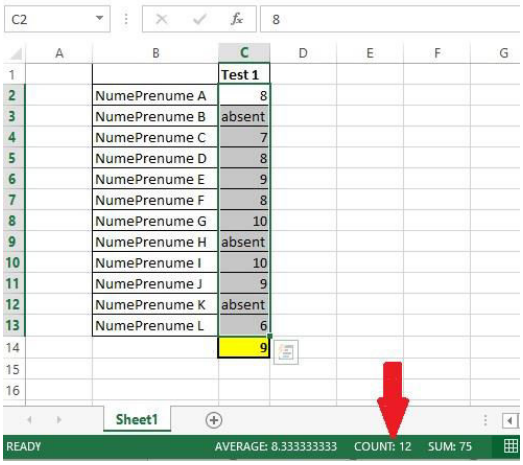
Argumentele

Value1…[value2] – numerele sau referințele care conțin valorile pe care dorim să le numărăm.

De exemplu: Să presupunem că avem în gestiune lista participanților la un curs cu rezultatele obținute în urma unui test. Nu toți participanții la curs au fost prezenți la test. Pentru a verifica numărul celor care au obținut rezultate la testul final, vom folosi funcția **=COUNT**. Aceasta ne oferă numărul celulelor care conțin numere. Dacă vom selecta coloana care conține rezultatele la test, putem observa că în foaia de calcul, în colțul din dreapta-jos, avem valoarea 12 la COUNT. Aceasta returnează numărul de celule selectate (în cazul exemplului nostru, sunt 12 participanți la curs, din care 9 au obținut rezultate la test).



	Test 1
NumePrenume A	8
NumePrenume B	absent
NumePrenume C	7
NumePrenume D	8
NumePrenume E	9
NumePrenume F	8
NumePrenume G	10
NumePrenume H	absent
NumePrenume I	10
NumePrenume J	9
NumePrenume K	absent
NumePrenume L	6
	9



	Test 1
NumePrenume A	8
NumePrenume B	absent
NumePrenume C	7
NumePrenume D	8
NumePrenume E	9
NumePrenume F	8
NumePrenume G	10
NumePrenume H	absent
NumePrenume I	10
NumePrenume J	9
NumePrenume K	absent
NumePrenume L	6
	9

Funcția

=COUNTA(value1, [value2], …)

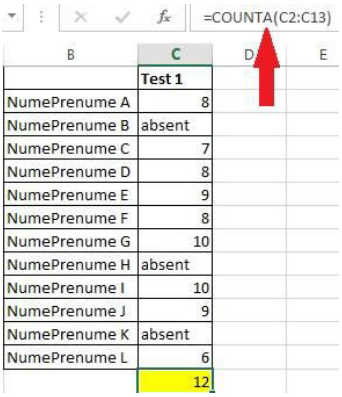
Descriere

Funcția **=COUNTA** (tot din categoria funcțiilor Statistice) returnează numărul de celule dintr-un șir care conțin o anumită informație (celule care nu sunt goale).

Argumentele

Value1…[value2] – numerele sau referințele care conțin valorile pe care dorim să le numărăm (celule care nu sunt goale).

De exemplu: Pornind de la datele din exemplul anterior, dacă vom folosi funcția **=COUNTA** rezultatul obținut va fi 12, deoarece funcția returnează numărul de celule care nu sunt goale.



	Test 1
NumePrenume A	8
NumePrenume B	absent
NumePrenume C	7
NumePrenume D	8
NumePrenume E	9
NumePrenume F	8
NumePrenume G	10
NumePrenume H	absent
NumePrenume I	10
NumePrenume J	9
NumePrenume K	absent
NumePrenume L	6
	12

Funcția

=COUNTBLANK(range)

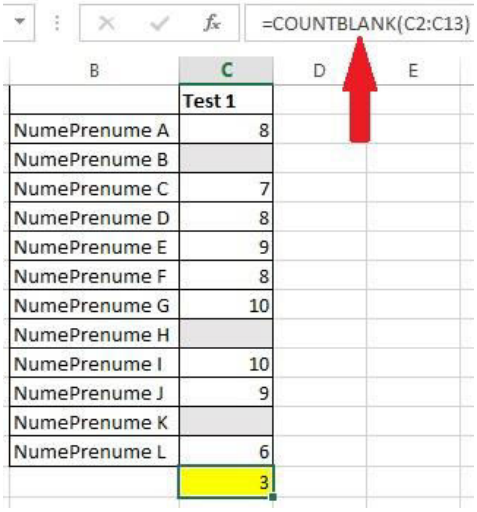
Descriere

Funcția **=COUNTBLANK** returnează numărul de celule dintr-un șir care NU conțin o anumită informație (celule care sunt goale).

Argumentele

Range – zona celulelor în care dorim să identificăm numărul de celule care nu conțin informații.

De exemplu: Tot în exemplul anterior, să presupunem că participanții care nu au dat testul nu au fost înregistrați ca "absent", ci nu au înregistrată o notă (celulele în cazul lor sunt goale). Dacă vom folosi funcția **=COUNTBLANK** rezultatul obținut va fi 3, deoarece funcția returnează numărul de celule care nu conțin nicio informație (sunt goale – *blank*).



Funcțiile

=MIN(number1, [number2], …)

=MAX(number1, [number2], …)

Descriere

Cele două funcții din categoria funcțiilor statistice returnează valoarea cea mai mică (**=MIN**), respectiv cea mai mare (**=MAX**), dintr-o serie de date. Funcția exclude din seria de date celulele care conțin text.

Argumentele

Number1…[numer2] – numerele sau referințele care conțin valorile pe baza cărora dorim să identificăm valoarea minimă, respectiv maximă.

De exemplu: Dacă într-o organizației avem înregistrați numărul de voluntari implicați în diferite activități, putem să calculăm folosind funcția **=MAX** numărul maxim de voluntari implicați, respectiv folosind funcția **=MIN** vom obține numărul cel mai mic de voluntari implicați în activitățile organizației.

A	B	C
nr. voluntari		
25	**80**	=MAX(A2:A100)
28	**1**	=MIN(A2:A100)
32		
30		
4		
58		
73		
58		
6		
42		
76		
26		
71		
56		
75		

Funcția

=SUBTOTAL(function_num, ref1, [ref2], …)

Descriere

Funcția **=SUBTOTAL** returnează rezultatele aplicării principalelor funcții (sumă, medie, numărare, minim, maxim, înmulțire) pe o listă de date. Aceste rezultate sunt de fapt subtotaluri, deoarece nu sunt luate în calcul toate datele, ci doar o parte din ele după filtrare (selecție pe categorii).

Argumentele

Function_num – codurile aferente funcției utilizate: 1 – AVERAGE, 2 – COUNT, 103 – COUNTA, 4 – MAX, 5 – MIN, 6 – PRODUCT, 7 – STDEV, 8 – STDEVP, 9 –SUM, 10 – VAR, 11 – VARP. Funcțiile ignoră valorile ascunse. Dacă sunt folosite coduri de la 101 la 111, sunt luate în calcul și valorile ascunse. Funcțiile sunt descrise în capitolele din carte.

Ref1, 2… – zona, sau zonele, de celule pentru care dorim să calculăm subtotalurile.

O modalitate mai simplă (care este prezentată mai jos) pentru realizarea unor calcule pe selecție de date (categorii rezultate în urma filtrării) este cea disponibilă accesând meniul **DATA >Subtotals**.

Pentru început vom aplica funcția **=SUBTOTAL**.

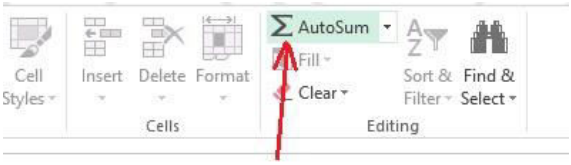
De exemplu: Să presupunem că avem o bază de date unde este înregistrat numărul de voluntari din 8 organizații care își desfășoară activitatea în trei orașe: Oradea, Arad și Timișoara. Putem să calculăm numărul total de voluntari fie utilizând funcția **=SUM**, fie utilizând funcția **=SUBTOTAL**. Se poate observa că rezultatele sunt aceleași dacă datele nu sunt filtrate (toate datele intră în calcul).

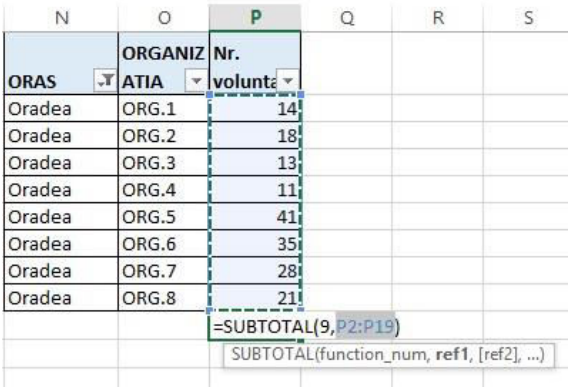
	A	B	C	D	E
1	ORAS	ORGANIZATIA	Nr. voluntari		
2	Oradea	ORG.1	14		
3	Arad	ORG.1	8		
4	Timișoara	ORG.1	10		
5	Oradea	ORG.2	18		
6	Timișoara	ORG.2	15		
7	Oradea	ORG.3	13		
8	Arad	ORG.2	12		
9	Oradea	ORG.4	11		
10	Timișoara	ORG.3	13		
11	Oradea	ORG.5	41		
12	Arad	ORG.3	11		
13	Timișoara	ORG.4	18		
14	Oradea	ORG.6	35		
15	Arad	ORG.4	10		
16	Oradea	ORG.7	28		
17	Timișoara	ORG.5	19		
18	Arad	ORG.5	13		
19	Oradea	ORG.8	21		
20					
21		TOTAL	310	=SUM(C2:C19)	
22		TOTAL/ORAS	310	=SUBTOTAL(9,C2:C19)	

În momentul în care aplicăm un filtru, de exemplu dorim să aflăm numărul de voluntari din Oradea, vom observa că rezultatul obținut când am folosit funcția **=SUM** nu se modifică. Indiferent de filtrarea datelor, funcția ia în considerare toate datele. În acest fel se întâmplă des să greșim și să raportăm date eronate. Valoarea corectă este cea obținută folosind funcția **=SUBTOTAL**. Această funcție include doar datele care corespund filtrului aplicat.

	A	B	C	D	E
1	ORAS	ORGANIZATIA	Nr. voluntari		
2	Oradea	ORG.1	14		
5	Oradea	ORG.2	18		
7	Oradea	ORG.3	13		
9	Oradea	ORG.4	11		
11	Oradea	ORG.5	41		
14	Oradea	ORG.6	35		
16	Oradea	ORG.7	28		
19	Oradea	ORG.8	21		
20					
21		TOTAL	310	=SUM(C2:C19)	
22		TOTAL/ORAS	181	=SUBTOTAL(9,C2:C19)	

Programul Excel intuiește corect când dorim să realizăm calcule pe o selecție de date. Astfel, dacă folosim filtrul inițial și calculăm suma direct accesând butonul pentru **AutoSum** ∑, funcția selectată automat nu va fi =**SUM**, ci =**SUBTOTAL**.

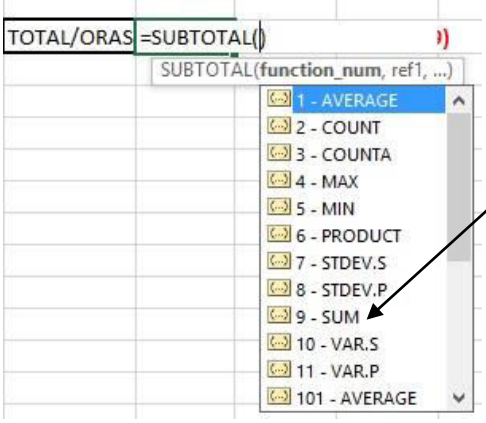




ORAS	ORGANIZATIA	Nr. volunta
Oradea	ORG.1	14
Oradea	ORG.2	18
Oradea	ORG.3	13
Oradea	ORG.4	11
Oradea	ORG.5	41
Oradea	ORG.6	35
Oradea	ORG.7	28
Oradea	ORG.8	21

Argumentele incluse în funcția =**SUBTOTAL** sunt: function_num, ref1, [ref2]. *Ref1…2* se referă la datele incluse în calcul (selecția datelor).

function_num – se referă la funcția utilizată pentru subtotal: AVERAGE, COUNT, COUNTA, MAX, MIN, PRODUCT, STDEV, SUM, VAR. Toate funcțiile au alocat un cod. Codurile de la 1 la 11 presupun includerea valorilor ascunse în calcul, iar codurile de la 101 la 111 exclud din funcție valorile ascunse. Funcțiile folosite pot fi selectate și direct cu cursorul.

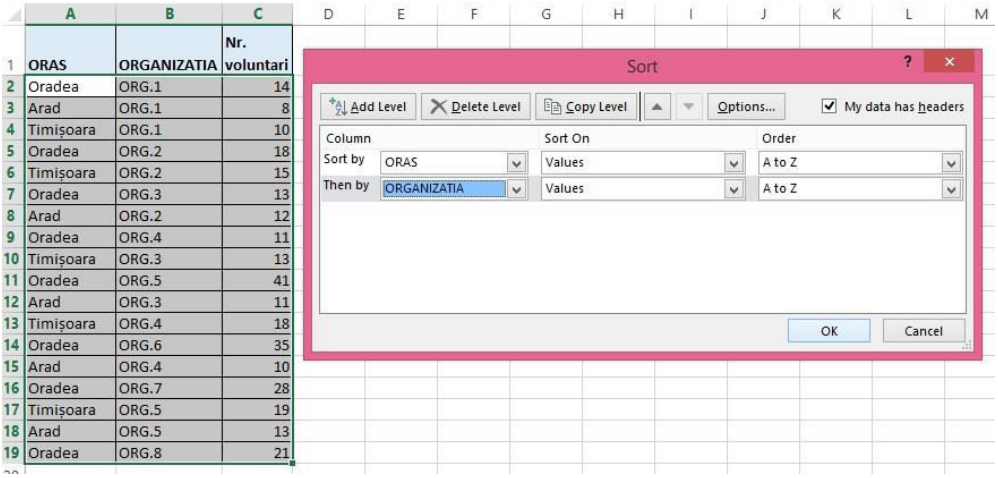


În exemplul de mai sus, am folosit codul 9. Astfel, am realizat o sumă, iar dacă baza de date ar fi avut rânduri ascunse (*hidden*) acestea ar fi fost incluse în calcul.

Realizarea Subtotalurilor accesând meniul DATA>Subtotal

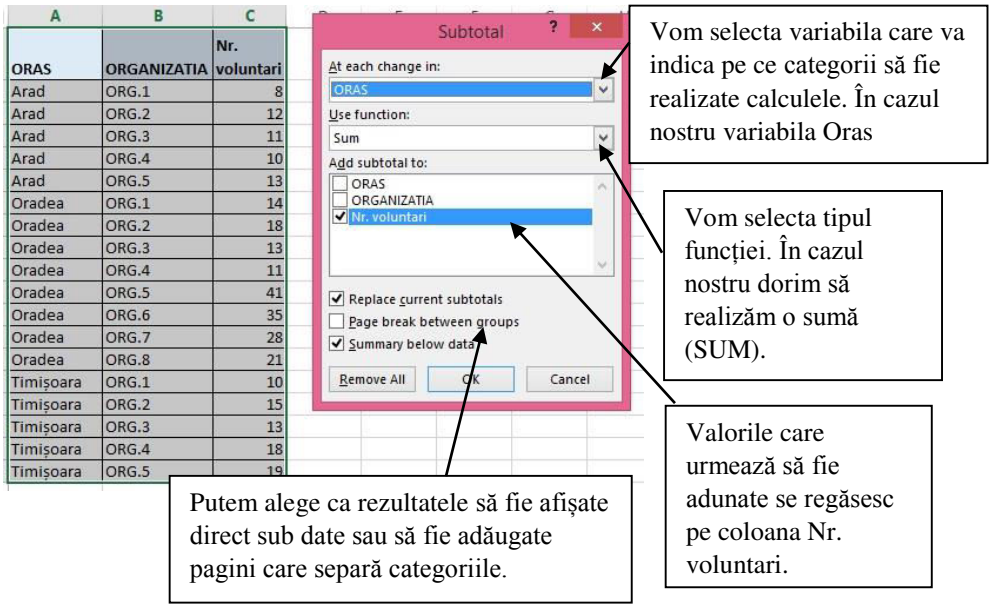
Pentru a aplica grupări ale datelor și realizarea unor subtotaluri în funcție de categoriile generate, este necesar pentru început să sortăm datele. Sortarea o vom face în funcție de variabilele care cuprind categoriile.

De exemplu: Vom utiliza datele din exemplul anterior. Vom sorta datele în funcție de variabila Oraș și variabila Organizație.



	ORAS	ORGANIZATIA	Nr. voluntari
2	Oradea	ORG.1	14
3	Arad	ORG.1	8
4	Timișoara	ORG.1	10
5	Oradea	ORG.2	18
6	Timișoara	ORG.2	15
7	Oradea	ORG.3	13
8	Arad	ORG.2	12
9	Oradea	ORG.4	11
10	Timișoara	ORG.3	13
11	Oradea	ORG.5	41
12	Arad	ORG.3	11
13	Timișoara	ORG.4	18
14	Oradea	ORG.6	35
15	Arad	ORG.4	10
16	Oradea	ORG.7	28
17	Timișoara	ORG.5	19
18	Arad	ORG.5	13
19	Oradea	ORG.8	21

Vom selecta oricare dintre celulele din tabel și vom accesa **DATA>Subtotal**.


Subtotal
At each change in:
ORAS
Use function:
Sum
Add subtotal to:
ORAS
ORGANIZATIA
Nr. voluntari
Replace current subtotals
Page break between groups
Summary below data
Remove All
OK
Cancel
Vom selecta variabila care va indica pe ce categorii să fie realizate calculele. În cazul nostru variabila Oras
Vom selecta tipul funcției. În cazul nostru dorim să realizăm o sumă (SUM).
Valorile care urmează să fie adunate se regăsesc pe coloana Nr. voluntari.
Putem alege ca rezultatele să fie afișate direct sub date sau să fie adăugate pagini care separă categoriile.

ORAS	ORGANIZATIA	Nr. voluntari
Arad	ORG.1	8
Arad	ORG.2	12
Arad	ORG.3	11
Arad	ORG.4	10
Arad	ORG.5	13
Oradea	ORG.1	14
Oradea	ORG.2	18
Oradea	ORG.3	13
Oradea	ORG.4	11
Oradea	ORG.5	41
Oradea	ORG.6	35
Oradea	ORG.7	28
Oradea	ORG.8	21
Timișoara	ORG.1	10
Timișoara	ORG.2	15
Timișoara	ORG.3	13
Timișoara	ORG.4	18
Timișoara	ORG.5	19

După ce confirmăm cu OK, rezultatele obținute vor arăta astfel: avem totalurile calculate pe fiecare oraș în parte și valoarea totală (*Grand Total*).

	A	B	C
1	ORAS	ORGANIZATIA	Nr. voluntari
2	Arad	ORG.1	8
3	Arad	ORG.2	12
4	Arad	ORG.3	11
5	Arad	ORG.4	10
6	Arad	ORG.5	13
7	**Arad Total**		54
8	Oradea	ORG.1	14
9	Oradea	ORG.2	18
10	Oradea	ORG.3	13
11	Oradea	ORG.4	11
12	Oradea	ORG.5	41
13	Oradea	ORG.6	35
14	Oradea	ORG.7	28
15	Oradea	ORG.8	21
16	**Oradea Total**		181
17	Timișoara	ORG.1	10
18	Timișoara	ORG.2	15
19	Timișoara	ORG.3	13
20	Timișoara	ORG.4	18
21	Timișoara	ORG.5	19
22	**Timișoara Total**		75
23	**Grand Total**		310

Folosind butoanele din stânga (cele care separă categoriile), putem să ascundem datele pe baza cărora s-au făcut subtotalurile.

	A	B	C
1	ORAS	ORGANIZATIA	Nr. voluntari
2	**Grand Total**		310
3	**Arad Total**		54
4	Arad	ORG.1	8
5	Arad	ORG.2	12
6	Arad	ORG.3	11
7	Arad	ORG.4	10
8	Arad	ORG.5	13
9	**Oradea Total**		181
18	**Timișoara Total**		75
19	Timișoara	ORG.1	10
20	Timișoara	ORG.2	15
21	Timișoara	ORG.3	13
22	Timișoara	ORG.4	18
23	Timișoara	ORG.5	19

	A	B	C
1	ORAS	ORGANIZATIA	Nr. voluntari
2	**Grand Total**		310
3	**Arad Total**		54
9	**Oradea Total**		181
18	**Timișoara Total**		75

4 FUNCȚIILE CONDIȚIONALE

COUNTIF, COUNTIFS, SUMIF, SUMIFS, AVERAGEIF, AVERAGEIFS

Funcțiile condiționale pornesc de la funcțiile simple de numărare, adunare și calculare a mediei, la care se adaugă una sau mai multe condiții.

Funcția NUMĂRĂ DACĂ

=COUNTIF(range, criteria)

Descriere

Funcția este folosită pentru numărare în funcție de un criteriu. Aceasta verifică informația din celulele selectate care pot fi evaluate ca ADEVĂRATE (TRUE) sau FALSE (FALSE) în funcție de un anumit criteriu și returnează numărul situațiilor evaluate ca ADEVĂRATE. Mai simplu spus: funcția numără câte celule corespund unui anumit criteriu.

Argumentele

range – zona de celule care va fi supusă condiționării

criteria – criteriul/condiția verificată; informația din fiecare celulă din zona selectată va fi verificată dacă corespunde criteriului.

Cum citim funcția

Numără DACĂ în celulele selectate informația este cea corespunzătoare criteriului.

De exemplu: Dacă avem o coloană cu date referitoare la numele localității și dorim să aflăm de câte ori apare o anumită localitate, vom folosi **=COUNTIF** specificând la criteriu numele localității.

	A	B	C
1	Oras	Tip institutie	
2	Brașov	privat	
3	Oradea	privat	
4	Arad	privat	
5	Oradea	privat	
6	Arad	public	
7	Sibiu	public	
8	Constanța	public	
9	București	privat	
10	Timișoara	privat	
11	Arad	privat	
12	Oradea	public	
13	Sibiu	privat	
14	Timișoara	public	
15	=COUNTIF(A2:A14, "Oradea")		
16	COUNTIF(range, criteria)		

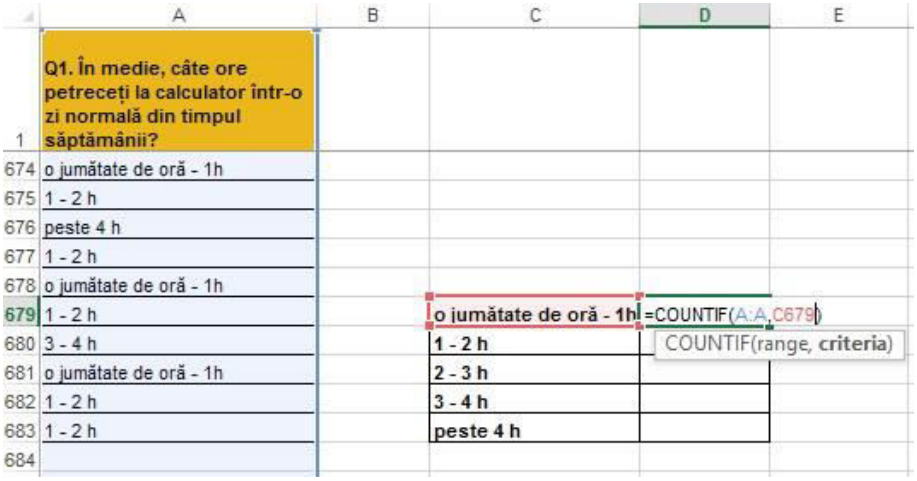
Rezultatul va fi numărul de câte ori apare "Oradea" (3), deoarece acesta este criteriul în funcție de care se face numărarea.

Un alt exemplu: În cadrul unui sondaj cu privire la competențele digitale am adresat următoarea întrebare: *În medie, câte ore petreceți la calculator într-o zi normală din timpul săptămânii?*, având variantele de răspuns: *o jumătate de oră - 1h, 1 - 2 h, 2 - 3 h, 3 - 4 h, peste 4 h*. În baza de date sunt înregistrate 682 de cazuri (am obținut răspunsurile la întrebare de la 682 de persoane). Dacă dorim să analizăm răspunsurile subiecților, un prim pas este să aflăm frecvențele înregistrate pentru fiecare categorie de răspuns. În acest sens trebuie să aflăm câți dintre cei 683 de subiecți au ales varianta de răspuns "o jumătate de oră - 1h", câți au răspuns "1-2h" ș.a.m.d.

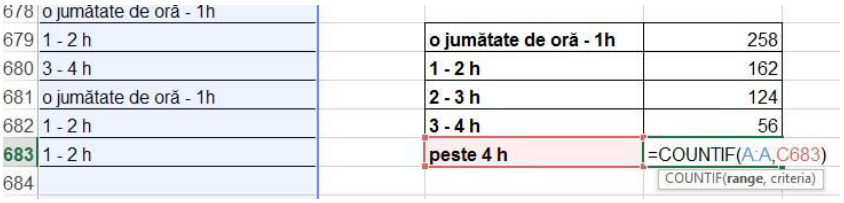
1	Q1. În medie, câte ore petreceți la calculator într-o zi normală din timpul săptămânii?			
662	1 - 2 h			
663	o jumătate de oră - 1h			
664	1 - 2 h			
665	o jumătate de oră - 1h			
666	1 - 2 h			
667	2 - 3 h			
668	o jumătate de oră - 1h			
669	o jumătate de oră - 1h			
670	o jumătate de oră - 1h			
671	o jumătate de oră - 1h			
672	2 - 3 h			
673	3 - 4 h			
674	o jumătate de oră - 1h			
675	1 - 2 h			
676	peste 4 h			
677	1 - 2 h			
678	o jumătate de oră - 1h			
679	1 - 2 h		o jumătate de oră - 1h	=COUNTIF(
680	3 - 4 h		1 - 2 h	COUNTIF(range, criteria)
681	o jumătate de oră - 1h		2 - 3 h	
682	1 - 2 h		3 - 4 h	
683	1 - 2 h		peste 4 h	

Vom utiliza funcția **=COUNTIF** pentru a număra de câte ori sunt menționate fiecare dintre variantele de răspuns. Pentru argumentul ***Range*** vom selecta toate datele

(toate cele 682 de celule care conțin răspunsurile). Deoarece consumăm prea mult timp pentru a selecta 682 de celule, am selectat direct coloana A care conține datele. Pentru argumentul ***Criteria*** vom selecta categoria de răspuns. Pentru a simplifica modul de lucru, am adăugat lista categoriilor de răspuns, astfel vom putea să selectăm direct criteriul din listă. Totodată, în acest fel putem să scriem funcția doar pentru prima categorie de răspuns ("o jumătate de oră - 1h"), iar apoi pentru celelalte categorii să multiplicăm funcția (folosim opțiunea auto-umplere).


A
B
C
D
E
1
Q1. În medie, câte ore petreceți la calculator într-o zi normală din timpul săptămânii?
674 o jumătate de oră - 1h
675 1 - 2 h
676 peste 4 h
677 1 - 2 h
678 o jumătate de oră - 1h
679 1 - 2 h
680 3 - 4 h
681 o jumătate de oră - 1h
682 1 - 2 h
683 1 - 2 h
684
o jumătate de oră - 1h
1 - 2 h
2 - 3 h
3 - 4 h
peste 4 h
=COUNTIF(A:A,C679)
COUNTIF(range, criteria)

678 o jumătate de oră - 1h
679 1 - 2 h
680 3 - 4 h
681 o jumătate de oră - 1h
682 1 - 2 h
683 1 - 2 h
684
o jumătate de oră - 1h 258
1 - 2 h 162
2 - 3 h 124
3 - 4 h 56
peste 4 h 82


678 o jumătate de oră - 1h
679 1 - 2 h
680 3 - 4 h
681 o jumătate de oră - 1h
682 1 - 2 h
683 1 - 2 h
684
o jumătate de oră - 1h 258
1 - 2 h 162
2 - 3 h 124
3 - 4 h 56
peste 4 h =COUNTIF(A:A,C683)
COUNTIF(range, criteria)

Funcția NUMĂRĂ DACĂ, folosind mai multe criterii

=COUNTIFS([criteria_range1,criteria1],[criteria_range2, criteria2], ….)

Descriere

Funcția este folosită pentru numărare în funcție de mai multe criterii. Aceasta verifică informația din celulele selectate care pot fi evaluate ca ADEVĂRATE sau FALSE în funcție criteriile menționate și returnează numărul situațiilor evaluate ca ADEVĂRATE. Mai simplu spus: funcția numără câte celule corespund tuturor criteriilor date.

Argumentele

criteria_range1…2 – zonele de celule care vor fi supuse condiționării

criteria1…2 – criteriile/condițiile verificate; informațiile din fiecare celulă din zonele selectate vor fi verificate dacă corespund criteriilor. Rezultatul va returna numărul situațiilor ADEVĂRATE după verificarea tuturor criteriilor.

Cum citim funcția

Numără DACĂ în celulele selectate informația este cea corespunzătoare criteriului 1 și numără DACĂ în celulele selectate informația este cea corespunzătoare unui alt criteriu etc..

De exemplu: Pornind de la exemplul anterior cu privire la sondajul realizat pe un lot de 682 de subiecți, să presupunem că datele au fost colectate în două orașe: Oradea și Cluj-Napoca. Dorim să aflăm cum se prezintă răspunsurile înregistrate la întrebarea *În medie, câte ore petreceți la calculator într-o zi normală din timpul săptămânii?* (analizată în exemplul anterior) în funcție de localitatea în care sunt subiecții.

Pentru început, folosind funcția =**COUNTIF** vom vedea că sunt 312 de subiecți din Oradea și 370 din Cluj-Napoca.

A
B
C
D
E
F
1
Q1. În medie, câte ore petreceți la calculator într-o zi normală din timpul săptămânii?
Localitate
674 o jumătate de oră - 1h Oradea
675 1 - 2 h Cluj-Napoca
676 peste 4 h Oradea
677 1 - 2 h Cluj-Napoca
678 o jumătate de oră - 1h Oradea
679 1 - 2 h Cluj-Napoca
680 3 - 4 h Oradea
681 o jumătate de oră - 1h Oradea
682 1 - 2 h Cluj-Napoca
683 1 - 2 h Oradea
684
685
Oradea 312
Cluj-Napoca =COUNTIF(B:B,D683)
COUNTIF(range, criteria)

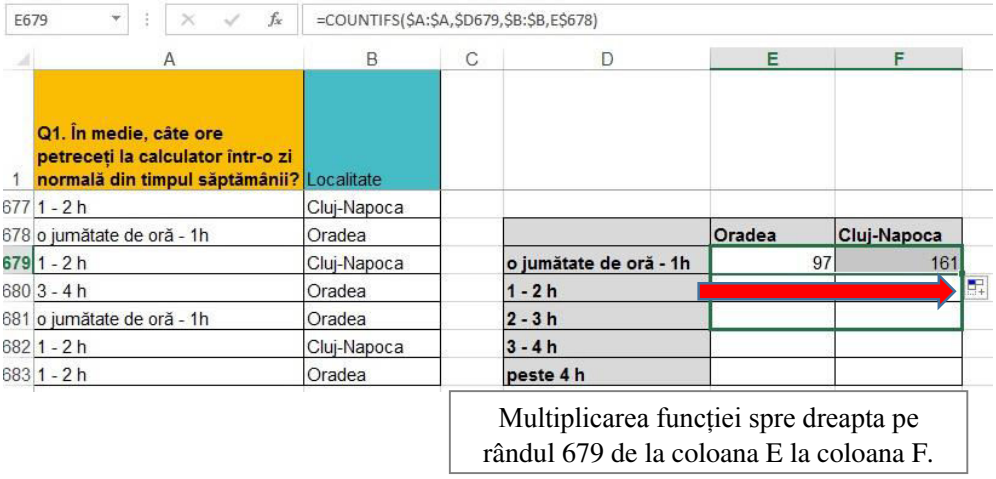
Apoi vom folosi funcția **=COUNTIFS** pentru a număra câte răspunsuri "o jumătate de oră – 1h" sunt în Oradea, respectiv în Cluj-Napoca, apoi câte răspunsuri "1-2h" sunt în Oradea, respectiv în Cluj-Napoca ș.a.m.d. Observăm că utilizăm doua criterii: categoria de răspuns și localitatea.

Deoarece dorim să multiplicăm funcția și pe coloane și pe rânduri, este necesar să utilizăm referințele mixte și absolute. În cazul argumentelor ***Range*** vom "bloca" coloanele: vom adăuga semnul $ în fața literei A, respectiv în fața literei B (care blochează coloanele în care sunt introduse datele pentru care dorim să aplicăm criteriile; în coloana A sunt categoriile de răspuns referioare la timpul petrecut la calculator; în coloana B sunt localitățile). Adăugarea semnului $ o putem face simplu apăsând tasta F4.

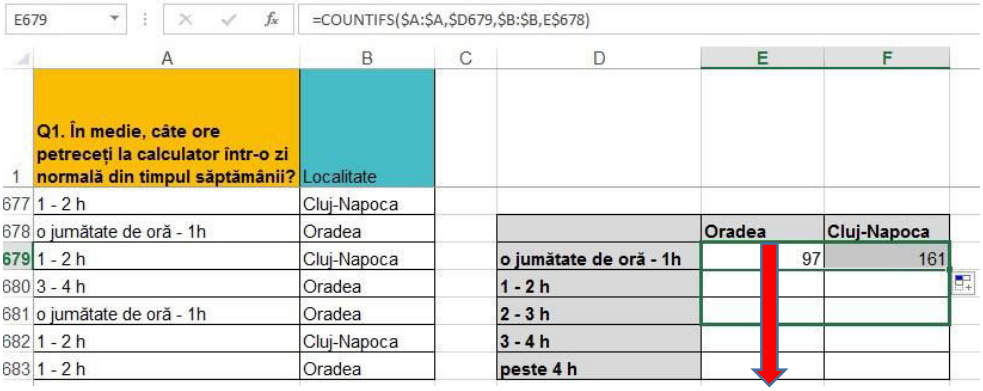
Pentru criterii (***Criteria 1*** și ***Criteria 2***) utilizăm referințele mixte. În funcție de situație, vom bloca fie coloana, fie rândul.

Pentru **Criteria 1**, care se referă la categoriile de răspuns ale subiecților, trebuie să permitem prin multiplicare în jos (pe coloană) programului Excel să selecteze fiecare dintre următoarele rânduri, dar la multiplicarea spre dreapta (pe rând) a funcției

(pentru coloana care conține datele aferente localității Cluj-Napoca), selecția trebuie să rămână blocată la referința inițială (coloana D) și să nu sară la următoarea coloană (E) din dreapta.



Multiplicarea funcției spre dreapta pe rândul 679 de la coloana E la coloana F.



Multiplicarea funcției în jos, pe coloana E, de la rândul 679 la rândurile 680, 681…

Referința pentru argumentul **Criteria 1** este celula D679. Când vom multiplica funcția în jos, pe coloana D, trebuie să îi permitem programului să utilizeze următoarele referințe (corespunzătoare următoarelor categorii de răspuns care sunt introduse pe rândurile următoare în jos). Astfel, nu vom adăuga semnul $ în fața rândului (679). Dar când vom multiplica funcția în dreapta, pe rândul 679, nu dorim referința să se deplaseze de pe D679 pe E679 (deoarece datele noastre sunt introduse în coloana D). Astfel, vom adăuga (vom bloca) semnul $ înaintea coloanei (D).

Referința pentru argumentul **Criteria 2** este celula E678 (care conține numele localității Oradea). Când vom multiplica funcția în jos, pe coloana D, nu trebuie să îi permitem programului să utilizeze referințele de pe rândurile următoare. Astfel,

vom adăuga semnul $ în fața rândului (678). Dar când vom multiplica funcția în dreapta, pe rând, dorim ca referința să se deplaseze de pe E678 pe F678 (unde se află introdusă localitatea Cluj-Napoca.

Rezultatele se prezintă astfel:

	A	B	C	D	E	F
1	**Q1. În medie, câte ore petreceți la calculator într-o zi normală din timpul săptămânii?**	Localitate				
677	1 - 2 h	Cluj-Napoca				
678	o jumătate de oră - 1h	Oradea			**Oradea**	**Cluj-Napoca**
679	1 - 2 h	Cluj-Napoca		**o jumătate de oră - 1h**	97	161
680	3 - 4 h	Oradea		**1 - 2 h**	113	49
681	o jumătate de oră - 1h	Oradea		**2 - 3 h**	58	66
682	1 - 2 h	Cluj-Napoca		**3 - 4 h**	22	34
683	1 - 2 h	Oradea		**peste 4 h**	22	60
684						

Funcțiile utilizate pentru fiecare celulă sunt următoarele:

	Oradea	Cluj-Napoca
o jumătate de oră - 1h	=COUNTIFS($A:$A,$D679, $B:$B,E$678)	=COUNTIFS($A:$A,$D679, $B:$B,F$678)
1 - 2 h	=COUNTIFS($A:$A,$D680, $B:$B,E$678)	=COUNTIFS($A:$A,$D680, $B:$B,F$678)
2 - 3 h	=COUNTIFS($A:$A,$D681, $B:$B,E$678)	=COUNTIFS($A:$A,$D681, $B:$B,F$678)
3 - 4 h	=COUNTIFS($A:$A,$D682, $B:$B,E$678)	=COUNTIFS($A:$A,$D682, $B:$B,F$678)
peste 4 h	=COUNTIFS($A:$A,$D683, $B:$B,E$678)	=COUNTIFS($A:$A,$D683, $B:$B,F$678)

Un alt exemplu: În cadrul unui sondaj am adresat subiecților (243 de persoane) o serie de întrebări cu privire la satisfacția față de diferite aspecte la locul de muncă: *Cât de mulțumit/ă sunteți cu privire la următoarele aspecte la locul actual de muncă? A. Salarizare; B. Siguranța locului de muncă; C. Programul de lucru; D. Relațiile cu colegii; E. Relația cu superiorii; F. Volumul de muncă; G. Oportunitățile privind promovarea; H. Oferta privind programele de instruire; I. Libertate de decizie*. Variantele de răspuns sunt: Foarte nemulțumit/ă, Nemulțumi/ă, Mulțumit/ă, Foarte mulțumit/ă. Pentru fiecare subiect a fost înregistrat şi sexul.

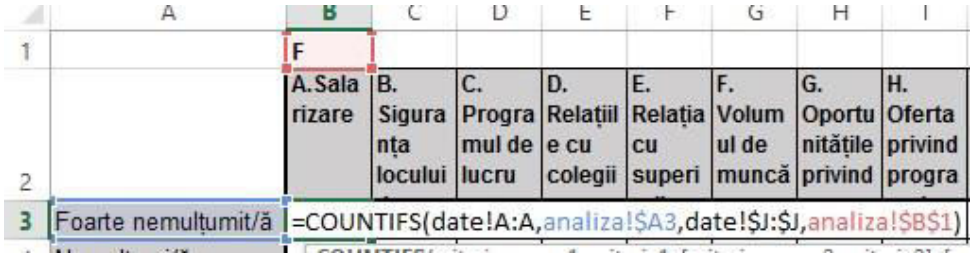
Răspunsurile le vom înregistra în baza de date:

	A	B	C	D	E	F	G	H	I	J
1	*Cât de mulțumit/ă sunteți cu privire la următoarele aspecte la locul actual de muncă?*									
2	A. Salarizare	B. Siguranța locului de muncă	C. Programul de lucru	D. Relațiile cu colegii	E. Relația cu superiorii	F. Volumul de muncă	G. Oportunitățile privind promovare	H. Oferta privind programele de	I. Libertate de decizie	GENUL
159	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Foarte mulțum	Mulțumit/ă	Mulțumit/ă	Nemulțumi/ă	Nemulțumi/ă	Nemulțumi/ă	F
160	Nemulțumi/ă	Mulțumit/ă	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Mulțumit/ă	Nemulțumi/ă	Mulțumit/ă	Mulțumit/ă	F
161	Mulțumit/ă	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțumit/ă	F
162	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Mulțumit/ă	Nemulțumi/ă	Nemulțumi/ă	Mulțumit/ă	M
163	Foarte mulțum	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțumit/ă	M
164	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Mulțumit/ă	Mulțumit/ă	Foarte nemulț	Foarte nemulțumit/ă	M
165	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțum	Foarte mulțumit/ă	F
166	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Mulțumit/ă	Mulțumit/ă	Nemulțumi/ă	Foarte mulțumit/ă	F
167	Foarte mulțum	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Foarte mulțum	Mulțumit/ă	Mulțumit/ă	F
168	Mulțumit/ă	Mulțumit/ă	Mulțumit/ă	Nemulțumi/ă	Mulțumit/ă	Nemulțumi/ă	Mulțumit/ă	Nemulțumi/ă	Mulțumit/ă	F
169	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Mulțumit/ă	Foarte mulțum	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Foarte mulțumit/ă	F
170	Mulțumit/ă	Foarte mulțum	Mulțumit/ă	Foarte mulțum	Foarte mulțum	Mulțumit/ă	Mulțumit/ă	Nemulțumi/ă	Foarte mulțumit/ă	F

Pentru a analiza răspunsurile obținute la aceste întrebări, vom folosi funcția **=COUNTIFS**. Dorim să aflăm câte femei au răspuns la prima întrebare că sunt Foarte mulțumite, apoi că sunt Mulțumite ș.a.m.d.., și răspunsurile date de bărbați la fiecare întrebare. Pentru a nu încărca foaia de calcul, vom insera o nouă foaie de calcul unde vom adăuga analizele. Prima foaie care cuprinde baza de date o vom denumi *Date*, iar a doua foaie de calcul o vom nota cu *Analize*.

Vom insera pentru început în foaia de calcul *Analize* lista categoriilor de răspuns (pentru a putea prin selecție să multiplicăm funcțiile) și lista întrebărilor pentru care facem analiza pentru ambele sexe (F și M).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		F									M								
2		A.Salarizare	B. Siguranța locului	C. Programul de lucru	D. Relațiile cu colegii	E. Relația cu superi	F. Volumul de muncă	G. Oportunitățile privind	H. Oferta privind progra	I. Libertate de decizie	A.Salarizare	B. Siguranța locului	C. Programul de lucru	D. Relațiile cu colegii	E. Relația cu superi	F. Volumul de muncă	G. Oportunitățile privind	H. Oferta privind progra	I. Libertate de decizie
3	Foarte nemulțumit/ă																		
4	Nemulțumi/ă																		
5	Mulțumit/ă																		
6	Foarte mulțumit/ă																		



Pentru a obține numărul femeilor care au răspuns *Foarte mulțumită* la prima întrebare referitoare la *Salarizare* vom selecta următoarele argumente pentru funcția **=COUNTIFS**: seria de date înregistrată în coloana A din foaia de calcul "date" pentru care vom aplica criteriul introdus în celula A3 din foaia de calcul "analiza".

Vom bloca coloana A deoarece nu dorim prin multiplicarea funcției selecția să se deplaseze pe coloana B, în schimb în fața rândului 3 nu vom adăuga semnul $ pentru a permite în momentul multiplicării funcției selecția să coboare pe rândurile următoare. Pentru al doilea criteriu vom selecta seria de date din coloana J din foaia de calcul "date" (coloana referioare la sexul subiecților). Pentru această serie vom aplica criteriul introdus în celula B1 din foaia de calcul "analiza". În cazul acestei celule vom bloca și coloana și rândul (vom adauga semnul $ și în fața literei și a cifrei) deoarece nu dorim selecția să se deplaseze de pe celula respectivă în momenul în care multiplicăm funcția. Folosind același mod de lucru vom realiza calculele și pentru subiecții de sex masculin.

Rezultatele se prezintă astfel:

	A	B	C	D	E	F	G	H	I	J	K	L
1		F									M	
2		A.Salarizare	B. Siguranța locului	C. Programul de lucru	D. Relațiile cu colegii	E. Relația cu superi	F. Volumul de muncă	G. Oportunitățile privind	H. Oferta privind progra	I. Libertate de decizie	A.Salarizare	B. Siguranța locului
3	Foarte nemulțumit/ă	2	4	1	0	1	4	7	7	3	1	5
4	Nemulțumi/ă	14	9	2	5	3	12	28	25	13	19	6
5	Mulțumit/ă	82	76	79	=COUNTIFS(date!D:D,analiza!$A5,date!$J:$J,analiza!$B$1)							
6	Foarte mulțumit/ă	24	33	40	COUNTIFS(criteria_range1, criteria1, [criteria_range2, criteria2], [crite							
7												

Funcția ÎNSUMEAZĂ DACĂ

=SUMTIF(range, criteria, [sum_range])

Descriere

Funcția este folosită pentru a aduna anumite valori selectate în funcție de un criteriu. Sunt selectate celulele în funcție de criteriul dat și sunt adunate numerele regăstite în celulele respective.

Argumentele

range – zona de celule care va fi supusă condiționării

criteria – criteriul/condiția verificată; informația din fiecare celulă din zona selectată va fi verificată dacă corespunde criteriului

sum_range – zona de celule care conțin numerele pe care dorim să le adunăm.

Cum citim funcția

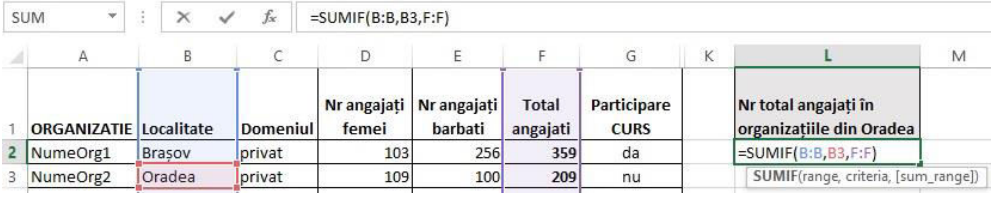
Însumează DACĂ în celulele selectate informația este cea corespunzătoare criteriului dat.

De exemplu: Avem în evidență un număr de 40 de organizații pentru care am înregistrat localitatea și tipul organizației, numărul de angajați (femei și bărbați),

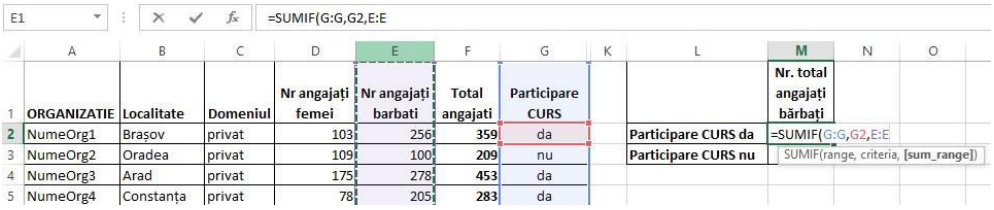
situația dacă angajații au participat la un curs de pregătire, respectiv numărul de participanți la curs (femei, bărbați)

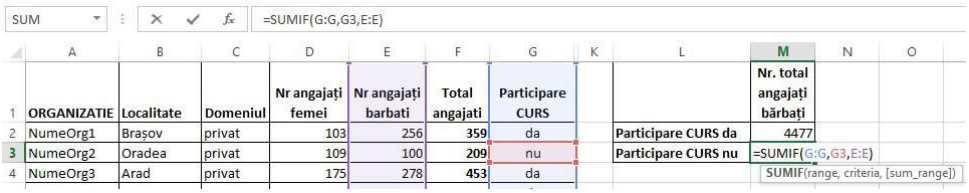
	A	B	C	D	E	F	G	H	I	J
1	ORGANIZATIE	Localitate	Domeniul	Nr angajați femei	Nr angajați barbati	Total angajati	Participare CURS	Nr participanți femei CURS	Nr participanți barbati CURS	TOTAL PARTICIPANTI CURS
2	NumeOrg1	Brașov	privat	103	256	**359**	da	56	53	109
3	NumeOrg2	Oradea	privat	109	100	**209**	nu	-	-	-
4	NumeOrg3	Arad	privat	175	278	**453**	da	45	36	81
5	NumeOrg4	Constanța	privat	78	205	**283**	da	31	36	67
6	NumeOrg5	Timișoara	privat	203	236	**439**	nu	-	-	-
7	NumeOrg6	Sibiu	public	146	192	**338**	nu	-	-	-
8	NumeOrg7	Brașov	privat	140	233	**373**	da	66	46	112

Dorim pentru început să aflăm câți angajați au în total organizațiile din Oradea. În acest sens, trebuie să selectăm celulele din coloana B care conțin informația "Oradea" și apoi să adunăm valorile care se regăsesc în coloana F. Vom folosi funcția **=SUMIF** pentru care vom adăuga următoarele argumente: pentru argumentul ***Range*** vom selecta coloana B (coloana în care se află datele din care vom selecta criteriul), pentru argumentul ***Criteria*** fie vom scrie direct criteriul "Oradea", fie vom selecta o celulă în care se află informația "Oradea", iar pentru argumentul [***Sum_range***] vom selecta coloana cu valorile pe care dorim să le însumăm (coloana F). Funcția o vom citi astfel: dacă pe coloana B informația este "Oradea", atunci însumează valorile are se regăsesc în coloana F.



Pe aceleași date, dorim să aflăm care este numărul total de angajați bărbați în cazul organizațiilor care au angajați care au participat la cursul de formare, respectiv pentru cei care nu au participat la curs. Folosind funcția **=SUMIF** vom adăuga argumentele: pentru ***Range*** vom selecta coloana G deoarece pe această coloană se regăsește criteriul, pentru ***Criteria*** vom selecta o celulă în care este introdus criteriul "da" (sau putem să îl scriem direct), iar pentru [***Sum_range***] vom selecta coloana E, deoarece aceasta cuprinde numărul de angajați bărbați pe care dorim să îl adunăm. În același mod vom calcula și numărul total de angajați bărbați în cazul organizațiilor ai căror angajați nu au participat la cursul de formare.





	A	B	C	D	E	F	G	K	L	M	N	O
1	ORGANIZATIE	Localitate	Domeniul	Nr angajați femei	Nr angajați barbati	Total angajati	Participare CURS			Nr. total angajați bărbați		
2	NumeOrg1	Brașov	privat	103	256	359	da		Participare CURS da	4477		
3	NumeOrg2	Oradea	privat	109	100	209	nu		Participare CURS nu	=SUMIF(G:G,G3,E:E)		
4	NumeOrg3	Arad	privat	175	278	453	da					

Funcția ÎNSUMEAZĂ DACĂ, folosind mai multe criterii

=SUMTIFS([sum_range],criteria_range1, criteria1, [criteria_range2, criteria2], …)

Descriere

Funcția este folosită pentru a aduna anumite valori selectate în funcție de mai multe criterii. Sunt selectate celulele care conțin valorile pe care dorim să le însumăm în funcție de mai multe criterii.

Argumentele

sum_range – zona de celule care conțin numerele pe care dorim să le adunăm

criteria_range1 – zona de celule care va fi supusă primei condiționări

criteria1 – criteriul/condiția verificată; informația din fiecare celulă din zona selectată va fi verificată dacă corespunde primului criteriu

criteria_range2 – zona de celule care va fi supusă celei de-a doua condiționări

criteria2 – criteriul/condiția verificată; informația din fiecare celulă din zona selectată va fi verificată dacă corespunde celui de-al doilea criteriu etc.

Pot fi selectate criterii multiple, în funcție de ceea ce dorim să analizăm.

Cum citim funcția

Însumează valorile selectate DACĂ în celulele selectate apoi informația este cea corespunzătoare primului criteriu, apoi în celulele selectate ulterior informația este cea corespunzătoare criteriului 2 s.a.m.d..

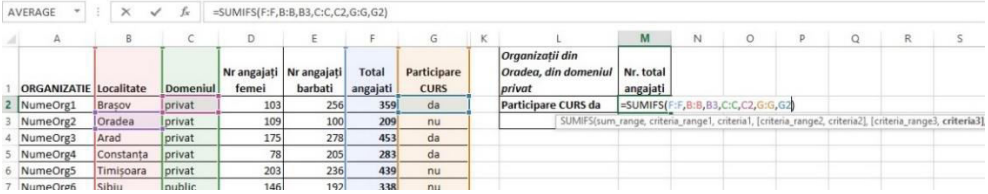
De exemplu: Pe aceeași bază de date privind cele 40 de organizații, dorim să aflăm numărul total de angajați bărbați în cazul organizațiilor din Oradea ai căror angajați au participat la cursul de formare. Față de exemplul prezentat anterior, avem un criteriu în plus, cel referitor la localitate. Vom avea primul criteriu: organizațiile să fie din Oradea, respectiv al doilea criteriu: angajații să fi participat la cursul de formare. Funcția folosită este **=SUMIFS** (având mai multe criterii). Vom selecta pentru început pentru argumentul ***Sum_range*** zona celulelor care conțin valorile pe care dorim să le adunăm (coloana E cu numărul de angajați bărbați), apoi vom selecta

celulele are conțin datele aferente primului criteriu (coloana B cu numele localităților), apoi vom selecta celula care conține criteriul în funcție de care se realizează selecția (putem să selectăm direct celula B3 care conține informația Oradea sau să scriem în glilimele "Oradea"), apoi vom selecta celulele care conțin datele aferente celui de-al doilea criteriu (coloana G care prezintă situația participarii la curs) și în final vom selecta celula care conține criteriul al doilea în funcție de care se face selecția celulelor (putem să selectăm direct celula G2 care conține informația "da" sau să scriem în ghilimele "da").

	ORGANIZATIE	Localitate	Domeniul	Nr angajați femei	Nr angajați barbati	Total angajati	Participare CURS
2	NumeOrg1	Brașov	privat	103	256	359	da
3	NumeOrg2	Oradea	privat	109	100	209	nu
4	NumeOrg3	Arad	privat	175	278	453	da
5	NumeOrg4	Constanța	privat	78	205	283	da
6	NumeOrg5	Timișoara	privat	203	236	439	nu
7	NumeOrg6	Sibiu	public	146	192	338	nu

	L	M
1	Organizații din Oradea	Nr. total angajați bărbați
2	Participare CURS da	=SUMIFS(E:E,B:B,B3,G:G,G2)

În continuare ne propunem să identificăm numărul total de angajați din organizațiile din Oradea din domeniul privat ai căror angajați au participat la cursul de formare profesională. În această situație avem 3 criterii de adăugat: organizația să fie din Oradea, organizația să fie din domeniul privat și organizația să fi participat la cursul de formare. Deoarece avem mai multe criterii, vom folosi funcția **=SUMIFS**. Vom selecta celulele care cuprind valorile pe care dorim să le însumăm (coloana F), apoi vom selecta celulele care cuprind datele din care vom selecta primul criteriu (coloana B), apoi vom selecta primul criteriu (informația din celula B3 "Oradea"), apoi vom selecta celulele care conțin datele aferente criteriului 2 (coloana C) și criteriul 2 (informația care se regăsește în celula C2 – "privat"), apoi vom selecta celulele care conțin datele aferente criteriului 3 (coloana G) și în final vom specifica criteriul 3 (vom selecta celula G2 - "da").



	ORGANIZATIE	Localitate	Domeniul	Nr angajați femei	Nr angajați barbati	Total angajati	Participare CURS
2	NumeOrg1	Brașov	privat	103	256	359	da
3	NumeOrg2	Oradea	privat	109	100	209	nu
4	NumeOrg3	Arad	privat	175	278	453	da
5	NumeOrg4	Constanța	privat	78	205	283	da
6	NumeOrg5	Timișoara	privat	203	236	439	nu
7	NumeOrg6	Sibiu	public	146	192	338	nu

	L	M
1	*Organizații din Oradea, din domeniul privat*	Nr. total angajați
2	Participare CURS da	=SUMIFS(F:F,B:B,B3,C:C,C2,G:G,G2)

Funcția MEDIA ARITMETICĂ DACĂ

=AVERAGEIF(range, criteria, [average_range])

Descriere

Funcția este folosită pentru a calcula media aritmetică a unor valori selectate în funcție de criteriu. Sunt selectate celulele în funcție de criteriul dat și este calculată media aritmetică a valorilor regăstite în celulele respective.

Argumentele

range – zona de celule care va fi supusă condiționării

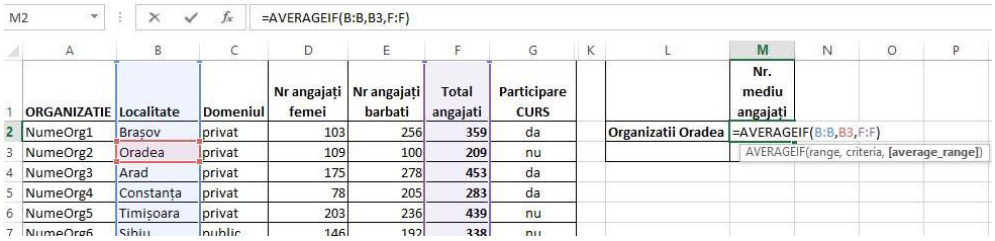
criteria – criteriul/condiția verificată; informația din fiecare celulă din zona selectată va fi verificată dacă corespunde criteriului

average_range – zona de celule care conțin numerele pe care dorim să le introducem în calcularea mediei aritmetice.

Cum citim funcția

Calculează media aritmetică DACĂ în celulele selectate informația este cea corespunzătoare criteriului dat.

De exemplu: Pe aceeași bază de date, dorim să calculăm numărul mediu de angajați din organizațiile din Oradea. Vom utiliza funcția **=AVERAGEIF** deoarece avem un singur criteriu. Pentru a adăuga argumentele funcției vom începe cu argumentul ***Range*** și vom selecta celulele care conțin datele aferente criteriului (coloana B care conține numele localităților), apoi vom selecta criteriul în funcție de care să fie selectate celulele (putem să selectăm direct celula B3 care conține informația "Oradea" sau să scriem în ghilimele "Oradea"). În final, pentru argumentul ***Average_range*** vom selecta celulele care conțin valorile pe baza cărora dorim să calculăm media aritmetică (coloana F care conține numărul de angajați).



	A	B	C	D	E	F	G	K	L	M	N	O	P
1	ORGANIZATIE	Localitate	Domeniul	Nr angajați femei	Nr angajați barbati	Total angajati	Participare CURS			Nr. mediu angajați			
2	NumeOrg1	Brașov	privat	103	256	359	da		Organizatii Oradea	=AVERAGEIF(B:B,B3,F:F)			
3	NumeOrg2	Oradea	privat	109	100	209	nu						
4	NumeOrg3	Arad	privat	175	278	453	da						
5	NumeOrg4	Constanța	privat	78	205	283	da						
6	NumeOrg5	Timișoara	privat	203	236	439	nu						
7	NumeOrg6	Sibiu	public	146	192	338	nu						

Funcția MEDIA ARITMETICĂ DACĂ, folosind mai multe criterii

=AVERAGEIFS([average_range],criteria_range 1, criteria1, [criteria_range2, criteria2], …)

Descriere

Funcția este folosită pentru a calcula media aritmetică a unor valori selectate în funcție de mai multe criterii. Sunt selectate celulele în funcție de criteriile date și este calculată media aritmetică a valorilor regăstite în celulele respective.

Argumentele

average_range – zona de celule care conțin valorile pe care dorim să le introducem în calcularea mediei aritmetice

Criteria_range1 – zona de celule care va fi supusă primei condiționări

Criteria1 – primul criteriu/condiția verificată; informația din fiecare celulă din zona selectată va fi verificată dacă corespunde criteriului

Criteria_range2 – zona de celule care va fi supusă celei de a doua condiționări

Criteria2 – al doilea criteriu/condiția verificată; informația din fiecare celulă din zona selectată va fi verificată dacă corespunde criteriului 2 etc.

Pot fi selectate criterii multiple, în funcție de ceea ce dorim să analizăm.

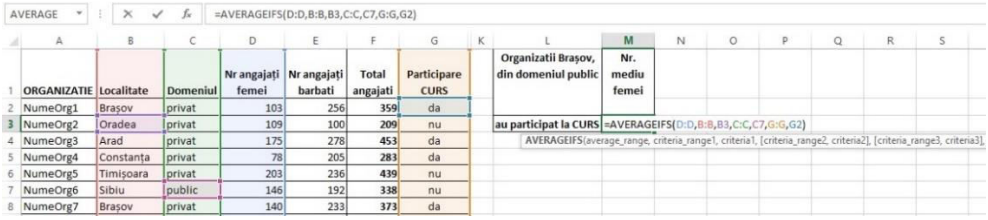
Cum citim funcția

Calculează media aritmetică DACĂ în celulele selectate informația este cea corespunzătoare primului criteriu dat, apoi selecția corespunde celui de-al doilea criteriu ș.a.m.d..

De exemplu: Pe datele din exemplul anterior, dorim să aflăm care este numărul mediu de angajați din organizațiile din Oradea, din domeniul public. Vom folosi funcția **=AVERAGEIFS**, având 2 criterii: organizațiile să fie din Oradea și organizațiile să fie din domeniul public. Argumentele funcției vor fi următoarele: pentru ***Average-range*** vom selecta celulele care conțin valorle pe care dorim să calculăm media aritmetică (coloana F care cuprinde numărul de angajați). Apoi vom selecta zona celulelor are conțin datele pe care le vom selecta potrivit primului criteriu (coloana B care cuprinde numele localităților) și vom adăuga primul criteriu (fie vom selecta direct celula B3 care conține informația "Oradea", fie vom scrie direct informația în ghilimele). În final, vom selecta celulele care cuprind datele aferente celui de-al doilea criteriu (coloana C cu privire la domeniul de activitate) și vom specifica criteriul al doi-lea (vom selecta celula C7 care conține informația "public" sau vom scrie noi informația).

AVERAGE
=AVERAGEIFS(F:F,B:B,B3,C:C,C7)
ORGANIZATIE Localitate Domeniul Nr angajați femei Nr angajați barbati Total angajati Participare CURS
NumeOrg1 Brașov privat 103 256 359 da
NumeOrg2 Oradea privat 109 100 209 nu
NumeOrg3 Arad privat 175 278 453 da
NumeOrg4 Constanța privat 78 205 283 da
NumeOrg5 Timișoara privat 203 236 439 nu
NumeOrg6 Sibiu public 146 192 338 nu
Organizatii Oradea, din domeniul public
Nr. mediu angajați
=AVERAGEIFS(F:F,B:B,B3,C:C,C7)
AVERAGEIFS(average_range, criteria_range1, criteria1, [criteria_range2, criteria2],

Dacă dorim să calculăm numărul mediu de angajați de sex feminin, din organizațiile din Brașov, din domeniul public, care au angajați care au participat la primul curs, vom folosi tot funcția **=AVERAGEIFS** deoarece avem mai multe criterii în funcție de care vom realiza calculul. Criteriile sunt: organizație din Oradea, din domeniul public, angajații organizației au participat la curs. Pentru adăugarea argumentelor funcției vom începe cu ***Average_range*** care reprezintă celulele care conțin valorile pe care dorim să le adăugăm în calcularea mediei, și anume numărul de angajați de sex feminin (coloana D). Apoi vom selecta ***Criteria_range1***, zona celulelor care cuprind datele din care vom selecta primul criteriu (coloana B referitoare la numele localității) și vom specifica primul criteriu (vom selecta celula B3 care conține "Oradea"). În pasul următor vom adăuga al doilea criteriu: vom selecta zona celulelor care cuprind datele aferente celui de-al doilea criteriu (coloana C referitoare la domeniu) și vom specifica cel de-al doilea criteriu (vom selecta celula C7 care conține informația "public"). În final vom adăuga cel de-al treilea criteriu: vom selecta zona celulelor care fac obiectul acestui criteriu (coloana G referitoare la participarea la curs) și vom specifica criteriul 3 (vom selecta celula G2 care conține informația "da").


AVERAGE
=AVERAGEIFS(D:D,B:B,B3,C:C,C7,G:G,G2)
ORGANIZATIE Localitate Domeniul Nr angajați femei Nr angajați barbati Total angajati Participare CURS
NumeOrg1 Brașov privat 103 256 359 da
NumeOrg2 Oradea privat 109 100 209 nu
NumeOrg3 Arad privat 175 278 453 da
NumeOrg4 Constanța privat 78 205 283 da
NumeOrg5 Timișoara privat 203 236 439 nu
NumeOrg6 Sibiu public 146 192 338 nu
NumeOrg7 Brașov privat 140 233 373 da
Organizatii Brașov, din domeniul public
Nr. mediu femei
au participat la CURS
=AVERAGEIFS(D:D,B:B,B3,C:C,C7,G:G,G2)
AVERAGEIFS(average_range, criteria_range1, criteria1, [criteria_range2, criteria2], [criteria_range3, criteria3],

5 FUNCȚIILE LOGICE

AND, OR, IF, NOT

Prezentarea funcțiilor logice

Funcțiile logice sunt utilizate pentru verificarea situațiilor în care datele corespund anumitor criterii (condiții logice). Acestea sunt cunoscute şi sub denumirea de funcții condiționale, deoarece probează îndeplinirea anumitor condiții. Aceste funcții permit elaborarea unor noi intrări (noi variabile) în baza de date, prin testarea condiționărilor. În acest sens, putem recodifica anumite valori ale unor variabile (text sau numerice) cu noi valori. În cadrul funcțiilor logice pot fi utilizați operatorii de comparație: egal cu (=), mai mic decât (<), mai mare decât (>), mai mic sau egal cu (<=), mai mare sau egal cu (>=) şi mai mic sau mai mare (< >). În sine aceşti operatori returnează expresii de tip ADEVĂRAT sau FALS pentru situația în care valoarea verificată corespunde, sau nu, criteriului folosit.

Principalele funcții logice sunt:

Funcția ȘI

=AND([logical1], [logical2], [logical3]…)

Descriere

Testează condiții care pot fi evaluate ca ADEVĂRATE (TRUE) sau FALSE (FALSE). Returnează valoarea ADEVĂRAT (TRUE) dacă toate condițiile logice corespund criteriului specificat.

Argumentele

Logical1, logical 2… - condițiile logice care sunt testate.

Cum citim funcția

ȘI prima condiție este îndeplinită, ȘI cea de-a doua condiție este îndeplinită, ȘI a treia condiție este îndeplinită etc.. Dacă toate condițiile sunt îndeplinite, atunci rezultatul funcției va fi TRUE. Dacă cel puțin una dintre condițiile verificate nu este îndeplinită, atunci rezultatul funcției va fi FALSE.

Funcția SAU

=OR(logical1], [logical2], [logical3]…)

Descriere

Testează condiții care pot fi evaluate ca ADEVĂRATE (TRUE) sau FALSE (FALSE). Returnează valoarea ADEVĂRAT (TRUE) dacă cel puțin una dintre condițiile logice corespunde criteriului specificat.

Argumentele

Logical1, logical 2… - condițiile logice care sunt testate.

Cum citim funcția

SAU prima condiție este îndeplinită, SAU cea de-a doua condiție este îndeplinită, SAU a treia condiție este îndeplinită etc.. Dacă cel puțin una dintre condiții este îndeplinită, atunci rezultatul funcției va fi TRUE. Dacă nici una dintre condițiile verificate nu este îndeplinită, atunci rezultatul funcției va fi FALSE.

Funcția DACĂ

=IF([logical_test], [value if true], [value if false])

Descriere

Testează dacă este îndeplinită o condiție logică și returnează o anumită valoare în situația în care expresia verificată este ADEVĂRATĂ, și o altă valoare dată expresia verificată este FALSĂ.

Argumente

Logical_test - condiția logică care este testată

Value if true – valoarea returnată dacă expresia logică este ADEVĂRATĂ

Value if false – valoarea returnată dacă expresia logică este FALSĂ.

Cum citim funcția

DACĂ expresia logică este adevărată, atunci returnează această valoare; dacă expresia este FALSĂ atunci returnează această valoare.

Funcția NU

=NOT(logical)

Descriere

Transformă expresiile logice identificate ca ADEVĂRATE (TRUE) în FALSE (FALSE) și invers.

Argumentele

Logical - condiția logică a cărei rezultat va fi inversat.

Cum citim funcția

Condiția logică testată NU nu este îndeplinită.

Utilizarea funcțiilor logice

De exemplu: Avem o bază de date care cuprinde rezultatele obținute de către angajații unei firme în urma participării la 3 teste de evaluare. Rezultatele la teste sunt înregistrate cu valori de la 10 la 100.

	A	B	C	D
1	NUME	scor TEST 1	scor TEST 2	scor TEST 3
2	Popescu	45	87	65
3	Șerbănescu	12	62	82
4	Florescu	56	85	64
5	Ionescu	85	74	92
6	Costinescu	12	45	73
7	Bădescu	45	25	18
8	Dupescu	65	55	82
9	Andronescu	24	69	49
10	Sorescu	85	37	44
11	Roșescu	61	51	74
12	Popescu	43	73	25

Pentru început dorim să identificăm angajații care au obținut valori mai mari sau egale cu 50 la toate cele trei teste. Vom realiza o nouă variabilă în baza de date, care va cuprinde expresie TRUE (ADEVĂRAT) dacă angajatul are rezultate mai mari sau egale la toate cele trei teste, respectiv FALSE (FALS) dacă angajatul are rezultatul mai mic de 50 cel puțin la unul dintre teste.

	A	B	C	D	E	F	G
1	NUME	scor TEST 1	scor TEST 2	scor TEST 3	situatie 1		
2	Popescu	45	87	65	=AND(B2>=50,C2>=50,D2>=50)		
3	Șerbănescu	12	62	82	FALSE		
4	Florescu	56	85	64	TRUE		
5	Ionescu	85	74	92	TRUE		
6	Costinescu	12	45	73	FALSE		
7	Bădescu	45	25	18	FALSE		
8	Dupescu	65	55	82	TRUE		
9	Andronescu	24	69	49	FALSE		
10	Sorescu	85	37	44	FALSE		
11	Rosescu	61	51	74	TRUE		

Condiția logică testată este: valoarea obținută la testul 1 ȘI cea de la testul 2 ȘI cea de la testul 3, sunt mai mari sau egale cu 50. Astfel, funcția folosită este **=AND**.

=AND(B2>=50,C2>=50,D2>=50)

Vom citi funcția în felul următor: ȘI valoarea din celula B2 corespunde condiției ">=50", ȘI valoarea din celula C2 corespunde condiției ">=50", ȘI valoarea din celula D2 corespunde condiției ">=50". Rezultatul funcției este TRUE dacă sunt îndeplinite toate cele 3 criterii, respectiv FALSE dacă cel puțin una dintre valorile din cele 3 celule nu corespunde criteriilor specificate.

În continuare, în baza de date dorim să realizăm o nouă variabilă care să țină cont de condiția: "valoare mai mare sau egală cu 50" pentru cel puțin unul dintre teste. Astfel, vom verifica dacă cel puțin una dintre valorile din cele 3 coloane este mai mare sau egală cu 50. Dacă această condiție este confirmată, în noua variabilă creeată vom avea expresia TRUE (ADEVĂRAT).

Condiția logică testată este: SAU valoarea obținută la testul 1, SAU cea de la testul 2, SAU cea de la testul 3, este mai mare sau egală cu 50. funcția folosită este **=OR**.

	A	B	C	D	E	F	G	H	I
1	NUME	scor TEST 1	scor TEST 2	scor TEST 3	situatie 1	situatie 2			
2	Popescu	45	87	65	FALSE	=OR(B2>=50,C2>=50, D2>=50)			
3	Șerbănescu	12	62	82	FALSE	OR(**logical1**, [logical2], [logical3], [logical4], ...)			
4	Florescu	56	85	64	TRUE	TRUE			
5	Ionescu	85	74	92	TRUE	TRUE			
6	Costinescu	12	45	73	FALSE	TRUE			
7	Bădescu	45	25	18	FALSE	FALSE			
8	Dupescu	65	55	82	TRUE	TRUE			
9	Andronescu	24	69	49	FALSE	TRUE			
10	Sorescu	85	37	44	FALSE	TRUE			

=OR(B2>=50,C2>=50, D2>=50)

Vom citi funcția în felul următor: SAU valoarea din celula B2 corespunde condiției ">=50", SAU valoarea din celula C2 corespunde condiției ">=50", SAU valoarea din celula D2 corespunde condiției ">=50". Rezultatul va fi sub forma TRUE dacă cel puțin una dintre condițiile testate este îndeplinită (angajatul a obținut valoare mai mare sau egală cu 50 măcar la unul dintre teste), respectiv FALSE dacă nici una dintre condițiile testate nu este îndeplinită (angajatul nu a obținut valori mai mari sau egale cu 50 la nici unul dintre teste).

Să presupunem că angajaților aflați în gestiunea noastră care au obținut valori mai mari de 60 la toate cele trei teste, dorim să le acordăm o mărire a salariului. În această situație, noua variabilă din baza de date va conține două tipuri de informații: "MĂRIRE SALAR" pentru angajații care au valori peste 60 la toate cele trei teste, respectiv "FĂRĂ MĂRIRE" pentru angajații care nu au obținut aceste valori la teste.

Condiția logică testată este: valoarea obținută la testul 1, ȘI cea de la testul 2, ȘI cea de la testul 3, sunt mai mari decât 60. Dacă aceste condiții sunt îndeplinite, și anume testul logic returnează valoarea ADEVĂRAT (TRUE), atunci rezultatul este "MĂRIRE SALAR", iar dacă nu sunt îndeplinite condițiile, și anume testul logic este FALS (FALSE), atunci rezultatul este "FĂRĂ MĂRIRE". În acest caz vom folosi funcția AND (ȘI) pentru a formula condițiile testate în funcția IF (DACĂ).

	A	B	C	D	E	F	G	H	I	J
1	NUME	scor TEST 1	scor TEST 2	scor TEST 3	situatie 3					
2	Popescu	45	87	65	=IF(AND(B2>60,C2>60,D2>60), "MARIRE SALAR","FARA MARIRE")					
3	Șerbănescu	12	62	82	FAR	AND(logical1, [logical2], [logical3], [logical4], ...)				
4	Florescu	56	85	64	FARA MARIRE					
5	Ionescu	85	74	92	MARIRE SALAR					
6	Costinescu	12	45	73	FARA MARIRE					
7	Bădescu	45	25	18	FARA MARIRE					
8	Dupescu	65	55	82	FARA MARIRE					
9	Andronescu	24	69	49	FARA MARIRE					

=IF(AND(B2>60, C2>60, D2>60), "MARIRE SALAR", "FARA MARIRE")

Vom citi funcția astfel: DACĂ, ȘI valoarea din celula B2 corespunde criteriului "mai mare decât 60", ȘI valoarea din celula C2 corespunde criteriului "mai mare decât 60", Și valoarea din celula D2 corespunde criteriului "mai mare decât 60", ATUNCI returnează rezultatul "MARIRE SALAR" (dacă expresia logică este adevărată), dacă expresia logică este falsă atunci returnează rezultatul "FARA MARIRE".

Să revenim la prima variabilă elaborată, cea care marchează situațiile în care angajații au obținut valori mai mari sau egale cu 50 la toate cele trei teste. Am folosit funcția **=AND**. Dacă dorim să verificăm situațiile în care angajații au obținut valori mai mici decât 50 la toate cele trei teste, nu este necesar să rescriem funcția =AND (B2<50, C2<50, D2<50), ci putem să folosim funcția **=NOT** pentru variabila

elaborată anterior. Funcția **=NOT** inversează argumentele ADEVĂRATE cu cele FALSE.

	A	B	C	D	E	F
1	NUME	scor TEST 1	scor TEST 2	scor TEST 3	situatie 1	valori <50
2	Popescu	45	87	65	FALSE	=NOT(E2)
3	Șerbănescu	12	62	82	FALSE	NOT(logical)
4	Florescu	56	85	64	TRUE	FALSE
5	Ionescu	85	74	92	TRUE	FALSE
6	Costinescu	12	45	73	FALSE	TRUE

Celula E2 conține funcția =AND(B2>=50,C2>=50,D2>=50). Dacă rezultatul din celula E2 este TRUE, rezultatul funcției =**NOT** din celula F2 va fi FALSE. Și invers: Dacă rezultatul din celula E2 este FALSE, rezultatul din celula F2 va fi TRUE.

Un alt exemplu: Să presupunem că avem în gestiune 100 de indivizi, pentru care avem înregistrate date cu privire la numărul de ani de experiență în muncă. Pornind de la această variabilă, dorim ca fiecare individ să îl clasificăm într-una dintre cele trei categorii: *Entry level* – Junior care include indivizii care au între 1 și 3 ani experiență de lucru, *Middle Level* pentru cei care au între 4 și 8 ani de experiență și *Senior level*, pentru indivizii care au peste 8 ani de experiență de muncă.

	A	B	C	D	E
1	INDIVID	nr. ani experienta	Entry level - Junior (1-3 ani)	Middle level (4-8 ani)	Senior level (peste 8 ani)
2	Nume 1	2			
3	Nume 2	3			
4	Nume 3	14			
5	Nume 4	5			
6	Nume 5	17			
7	Nume 6	18			
8	Nume 7	8			
9	Nume 8	2			
10	Nume 9	6			
11	Nume 10	16			
12	Nume 11	1			
13	Nume 12	6			
14	Nume 13	15			
15	Nume 14	15			
16	Nume 15	4			
17	Nume 16	17			
18	Nume 17	12			
19	Nume 18	5			

Pentru prima categorie (Coloana C), condiția testată este ca numărul de ani de experiență să fie între 1 și 3. DACĂ această condiție este îndeplinită (este adevărată), atunci individul respectiv va fi inclus în categoria "Junior". Indivizii care nu îndeplinesc această condiție, nu vor fi incluși într-o categorie anume (în dreptul lor nu va fi specificat nimic).

	A	B	C	D	
1	INDIVID	nr. ani experienta	Entry level - Junior (1-3 ani)	Middle level (4-8 ani)	Senio (peste
2	Nume 1	2	=IF(AND(B2>=1,B2<=3), "junior", "")		
3	Nume 2	3	junior		
4	Nume 3	14			
5	Nume 4	5			
6	Nume 5	17			

=IF(AND(B2>=1,B2<=3), "junior", "")

Vom citi funcția astfel: DACĂ valoarea din celula B2 este mai mare sau egală cu 1 ȘI este mai mică sau egală cu 3 (și anume este între 1 și 3), atunci categoria este "Junior". Dacă valoarea din celula B2 nu corespunde acestei condiții, atunci rezultatul funcției este un spațiu gol (între ghilimele nu vom adăuga nimic: ""). Cel de-al treilea argument din funcția =IF poate să nu mai fie menționat deloc în această situație.

Folosind același principiu (aceleași funcții), vom verifica pentru fiecare individ dacă va fi inclus în categoriile "Middle", respectiv "Senior".

	A	B	C	D	E	F
1	INDIVID	nr. ani experienta	Entry level - Junior (1-3 ani)	Middle level (4-8 ani)	Senior level (peste 8 ani)	
2	Nume 1	2	junior	=IF(AND(B2>=4,B2<=8), "middle","")		
3	Nume 2	3	junior			
4	Nume 3	14			senior	
5	Nume 4	5		middle		
6	Nume 5	17			senior	
7	Nume 6	18			senior	
8	Nume 7	8		middle		
9	Nume 8	2	iunior			

=IF(AND(B2>=4,B2<=8), "middle","")

Funcția va fi interpretată astfel: DACĂ valoarea din celula B2 este mai mare sau egală cu 4 ȘI este mai mică sau egală cu 8 (este între 4 și 8), atunci categoria este "Middle". Dacă valoarea din celula B2 nu corespunde acestei condiții, atunci rezultatul funcției este un spațiu gol (între ghilimele nu vom adăuga nimic: ""). Sau din nou putem lăsa necompletat argumentul acesta din funcție.

	A	B	C	D	E	F
1	INDIVID	nr. ani experienta	Entry level - Junior (1-3 ani)	Middle level (4-8 ani)	Senior level (peste 8 ani)	
2	Nume 1	2	junior		=IF(B2>8,"senior", "")	
3	Nume 2	3	junior			
4	Nume 3	14			senior	
5	Nume 4	5		middle		

=IF(B2>8,"senior", "")

Vom citi funcția astfel: DACĂ valoarea din celula B2 este mai mare decât 8, atunci categoria este "Senior". Dacă valoarea din celula B2 nu corespunde acestei condiții, atunci rezultatul funcției este un spațiu gol.

Cele 3 coloane care cuprind categoriile *Junior*, *Middle* și *Senior* le putem îmbina apoi într-o singura coloană. Vom avea astfel o nouă variabilă în baza de date.

F2 =CONCATENATE(C2,D2,E2)

	A	B	F	G	H	I
1	INDIVID	nr. ani experienta	LEVEL			
2	Nume 1	2	*junior*			
3	Nume 2	3	*junior*			
4	Nume 3	14	*senior*			
5	Nume 4	5	*middle*			
6	Nume 5	17	*senior*			
7	Nume 6	18	*senior*			
8	Nume 7	8	*middle*			

Pentru unirea celor trei coloane, cea mai potrivită funcție este **=CONCATENATE**([text1], [text2], [text3],…), care permite alăturarea mai multor șiruri de text într-un singur șir. Detalii privind utilizarea acestei funcții sunt prezentate în secțiunea dedicată funcțiilor pentru text.

În exemplul nostru, vom uni textele incluse în celulele C2, D2 și E2.

6 STATISTICA DESCRIPTIVĂ A DATELOR

FREQUENCY, MODE, MEDIAN, QUARTILE, SQRT, VAR, STDEV, CONFIDENCE.NORM

În cadrul acestui capitol sunt prezentate aspecte privind analiza descriptivă a datelor: realizarea tabelelor de frecvență și calcularea principalilor indicatori statistici. Materialul nu își propune definirea conceptelor utilizate în statistică și a indicatorilor statistici, ci prezentarea modalităților de calcul. În acest sens se recomandă ca cititorii să dețină cunoștințe minime privind statistica descriptivă.

Statistica descriptivă a datelor presupune în primul rând realizarea tabelelor de frecvență. În acest sens, pe baza datelor culese la nivel de eșantion sau la nivelul întregii populații, sunt identificate valorile variabilelor (notate în statistică cu *x*) și frecvențele absolute (notate cu *f*). Pornind de la aceste valori, sunt calculate frecvențele relative și cele cumulate. În funcție de tipul variabilelor utilizate, analiza descriptivă a datelor continuă cu principalii indicatori statistici.

Realizarea tabelelor de frecvență

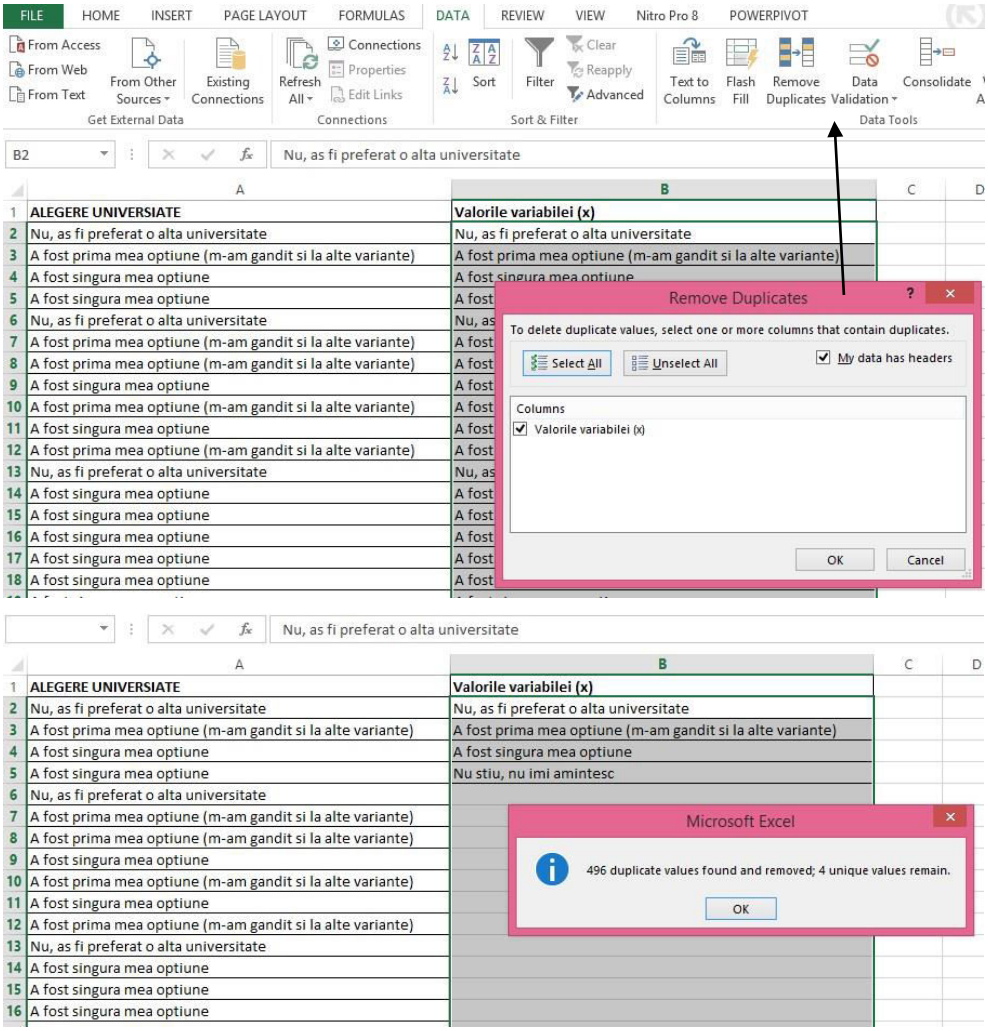
Pentru orice analiză statistică a datelor, trebuie să ținem cont de nivelurile de măsurare: nominal și ordinal (variabile calitative), de interval și de raport (variabile cantitative). Pentru realizarea tabelelor de frecvență un prim pas îl presupune identificarea frecvențelor absolute, iar apoi analiza descriptivă va continua cu realizarea frecvențelor relative și a celor cumulate. Pentru calcularea frecvențelor absolute, în cazul variabilelor calitative vom utiliza de regulă funcțiile condiționale (=**COUNTIF**). În ceea ce privește variabilele cantitative (numerice), putem să utilizăm de asemenea funcțiile condiționale sau funcția =**FREQUENCY**.

Tabelele de frecvență pentru variabilele calitative

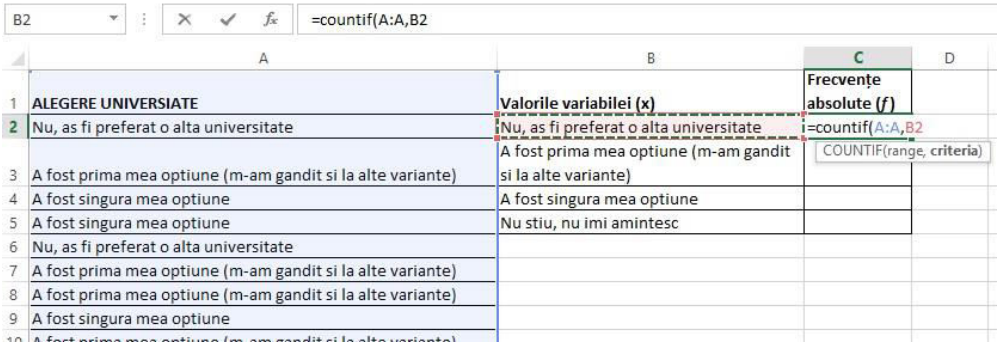
De exemplu: În cadrul unui sondaj realizat pe o populație de studenți am adresat următoarea întrebare subiecților: *Universitatea la care înveți a fost prima ta opțiune?*, având următoarele variante de răspuns: "*A fost singura mea opțiune*", "*A fost prima mea opțiune (m-am gândit și la alte variante)*", "*Nu, aș fi preferat o altă universitate*", "*Nu știu, nu îmi amintesc*". În baza de date au fost înregistrate 500 de cazuri (răspunsuri obținute de la studenți).

Pentru a identifica frecvențele absolute (*f*) vom utiliza funcția **=COUNTIF**, care permite numărarea în funcție de anumite criterii. Criteriile în funcție de care vom

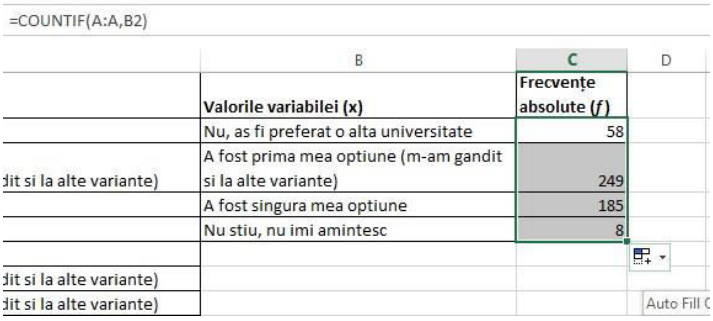
face numărarea sunt tocmai valorile variabilei (x). Pentru a le identifica, putem să folosim opțiunea ***Remove Duplicates*** accesând meniul **DATA**. Vom copia variabila pe o nouă coloană (sau într-o altă foaie de calcul) pentru a păstra baza de date inițială, iar pe selecția datelor copiate vom exclude din listă duplicatele. Astfel vom obține lista valorilor variabilei.



Pentru a calcula frecvențele absolute vom utiliza funcția **=COUNTIF**, selectând pentru argumentul funcției ***Range*** celulele în care sunt introduse răspunsurile obținute de la studenți (pe care dorim să realizăm numărarea; coloana A), iar pentru ***Criteria***, vom selecta valorile variabilei identificate anterior (prima valoare este introdusă în celula B2).



Pentru frecvențele absolute aferente următoarelor valori din listă ale variabilei, vom folosi opțiunea auto-umplere.

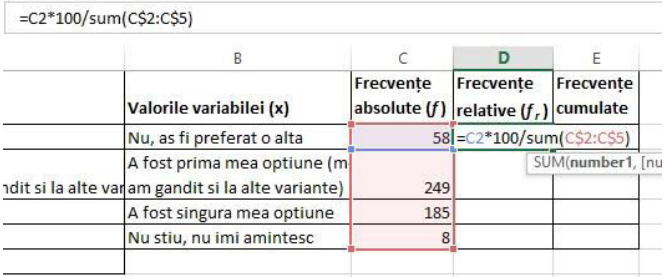


Pentru a calcula frecvențele relative, vom avea formula:

$$fr_i = \frac{f_i}{\sum f_i} * 100$$

unde f_i este frecvența absolută pentru fiecare valoare a variabilei și $\sum f_i = N$ reprezintă numărul total (numărul total de cazuri).

În foaia de calcul vom utiliza referințele: celula C2 conține frecvența absolută pentru prima valoare a variabilei, iar pentru numărul total N vom utiliza funcția **= SUM.** În cazul argumentelor pentru funcția =**SUM**, vom utiliza referințe relative (mixte) deoarece pentru calcularea următoarelor frecvențe relative din listă vom utiliza opțiunea auto-umplere. Multiplicând funcția în jos pe rânduri, dorim să păstrăm selecția referințelor pentru funcția =**SUM**. În acest sens, vom "bloca" selecția celulelor C2:C5, adăugând semnul $ înaintea cifrelor care notează rândul (C$2:C$5).



=C2*100/SUM(C$2:C$5)

	B	C	D	E
	Valorile variabilei (x)	**Frecvențe absolute (f)**	**Frecvențe relative (f_r)**	**Frecvențe cumulate**
	Nu, as fi preferat o alta	58	11.6	
ndit si la alte var	A fost prima mea optiune (m-am gandit si la alte variante)	249	49.8	
	A fost singura mea optiune	185	37	
	Nu stiu, nu imi amintesc	8	1.6	

ndit si la alte variante)

Frecvențele relative le vom interpreta astfel: 11.6% din cei 500 de subiecți declară că ar fi preferat o altă universitate decât cea la care studiază.

Frecvența cumulată se calculează prin însumarea frecvenței valorii variabilei și frecvențele aferente categoriilor anterioare.

Pentru prima valoare a variabilei, vom copia frecvența aferentă acesteia (deoarece nu există frecvențe anterioare ei). Frecvențele cumulate pentru valorile următoare ale variabilei le vom calcula adunând frecvența categoriei respective și frecvențele anterioare.

D2 =D2

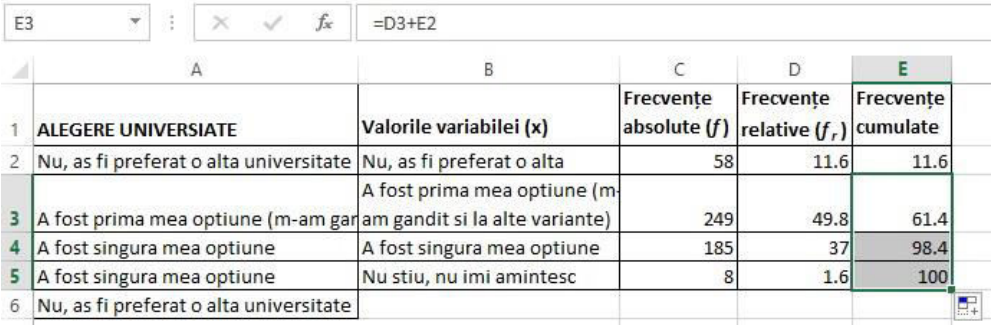
	A	B	C	D	E
1	**ALEGERE UNIVERSIATE**	**Valorile variabilei (x)**	**Frecvențe absolute (f)**	**Frecvențe relative (f_r)**	**Frecvențe cumulate**
2	Nu, as fi preferat o alta universitate	Nu, as fi preferat o alta	58	11.6	=D2
3	A fost prima mea optiune (m-am gar	A fost prima mea optiune (m-am gandit si la alte variante)	249	49.8	
4	A fost singura mea optiune	A fost singura mea optiune	185	37	
5	A fost singura mea optiune	Nu stiu, nu imi amintesc	8	1.6	

E2 =D3+E2

	A	B	C	D	E
1	**ALEGERE UNIVERSIATE**	**Valorile variabilei (x)**	**Frecvențe absolute (f)**	**Frecvențe relative (f_r)**	**Frecvențe cumulate**
2	Nu, as fi preferat o alta universitate	Nu, as fi preferat o alta	58	11.6	11.6
3	A fost prima mea optiune (m-am gar	A fost prima mea optiune (m-am gandit si la alte variante)	249	49.8	=D3+E2
4	A fost singura mea optiune	A fost singura mea optiune	185	37	
5	A fost singura mea optiune	Nu stiu, nu imi amintesc	8	1.6	
6	Nu, as fi preferat o alta universitate				

	A	B	C	D	E
1	ALEGERE UNIVERSIATE	Valorile variabilei (x)	Frecvențe absolute (f)	Frecvențe relative (f_r)	Frecvențe cumulate
2	Nu, as fi preferat o alta universitate	Nu, as fi preferat o alta	58	11.6	11.6
3	A fost prima mea optiune (m-am gar	A fost prima mea optiune (m-am gandit si la alte variante)	249	49.8	61.4
4	A fost singura mea optiune	A fost singura mea optiune	185	37	=D4+E3
5	A fost singura mea optiune	Nu stiu, nu imi amintesc	8	1.6	

Putem utiliza și de această dată opțiunea auto-umplere pentru a multiplica formula în jos pe rândurile aferente tuturor valorilor variabilei.

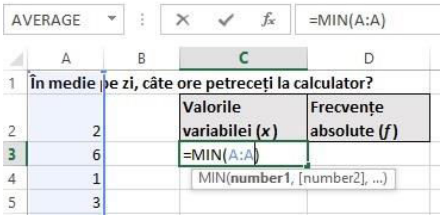


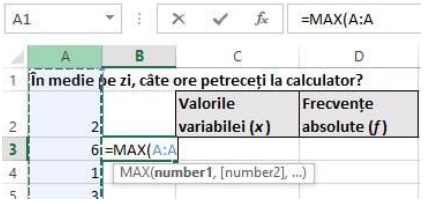
	A	B	C	D	E
1	ALEGERE UNIVERSIATE	Valorile variabilei (x)	Frecvențe absolute (f)	Frecvențe relative (f_r)	Frecvențe cumulate
2	Nu, as fi preferat o alta universitate	Nu, as fi preferat o alta	58	11.6	11.6
3	A fost prima mea optiune (m-am gar	A fost prima mea optiune (m-am gandit si la alte variante)	249	49.8	61.4
4	A fost singura mea optiune	A fost singura mea optiune	185	37	98.4
5	A fost singura mea optiune	Nu stiu, nu imi amintesc	8	1.6	100
6	Nu, as fi preferat o alta universitate				

Tabelele de frecvență pentru variabilele cantitative (numerice)

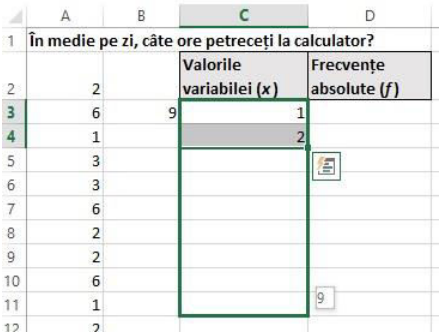
De exemplu: În cadrul unui sondaj realizat pe un eșantion de 465 de subiecți, am înregistrat răspunsurile la întrebarea: *În medie, câte ore petreceți pe zi la calculator?*. Am introdus răspunsurile în baza de date și dorim să realizăm tabelul de frecvență. Pentru variabilele numerice în mod uzual nu realizăm tabelele de frecvență, ci sunt utilizați principalii indicatori statistici pentru descrierea variabilei. Totuși, în această secțiune dorim să prezentăm modul de lucru și în cazul variabilelor numerice, folosind funcția = **FREQUENCY**.

Pentru început trebuie să identificăm valorile variabilei (x). Acestea sunt variantele de răspuns înregistrate. Deoarece sunt destul de multe cazuri în baza de date, sunt greu de urmărit toate valorile. Putem să optăm pentru modul de lucru prezentat anterior, și anume utilizarea opțiunii ***Remove duplicates*** (din meniul **DATA**) sau putem să identificăm valoarea minimă (folosind funcția =**MIN**), respectiv valoarea maximă (folosind funcția =**MAX**), iar cu opțiunea auto-umplere să adăugăm valorile intermediare (inserăm seria de date automat). Vom prezenta în cele ce urmează a doua variantă de lucru.





Am identificat valoarea minimă (1) și valoarea maximă (9) ale variabilei, iar apoi vom adăuga seria de date, de la 1 până la 9.



Pentru a calcula frecvențele absolute pentru fiecare valoare a variabilei vom utiliza funcția =**FREQUENCY**.

Funcția

= FREQUENCY (data_array, bins_array)

Descriere

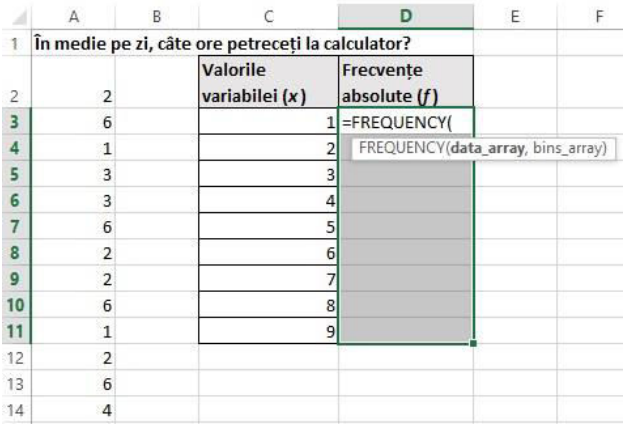
Funcția este folosită pentru calcularea frecvențelor absolute pe o serie de date. Calculează de câte ori apare o valoare într-o serie de date.

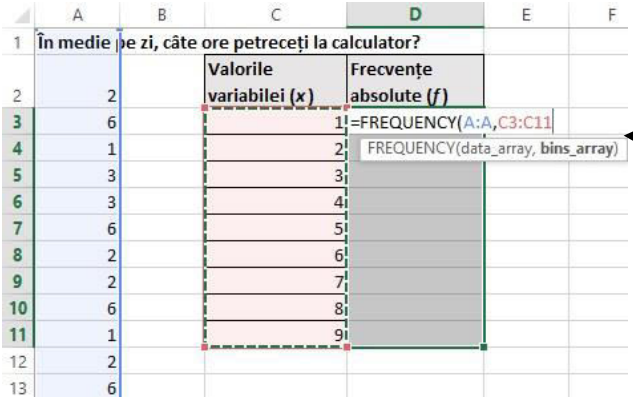
Argumentele

Data_array – zona de celule care cuprinde seria de date.

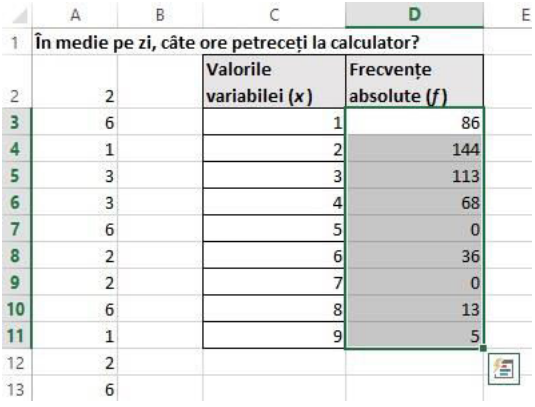
Bins_array – lista valorilor pe care dorim să le calculăm de câte ori apar în seria de date.

Selectăm pentru început zona celulelor în care dorim să calculăm frecvențele absolute, apoi (fără să renunțăm la selecție) scriem funcția =**FREQUENCY**. Pentru argumentul ***Data_array*** vom selecta seria de date (coloana A în care am înregistrat răspunsurile subiecților), iar pentru argumentul ***Bins_array*** vom selecta lista valorilor (valorile variabilei de la 1 la 9 introduse în coloana C). Apoi, pentru a calcula, apăsăm tastele **CTRL+SHIFT+ENTER**.





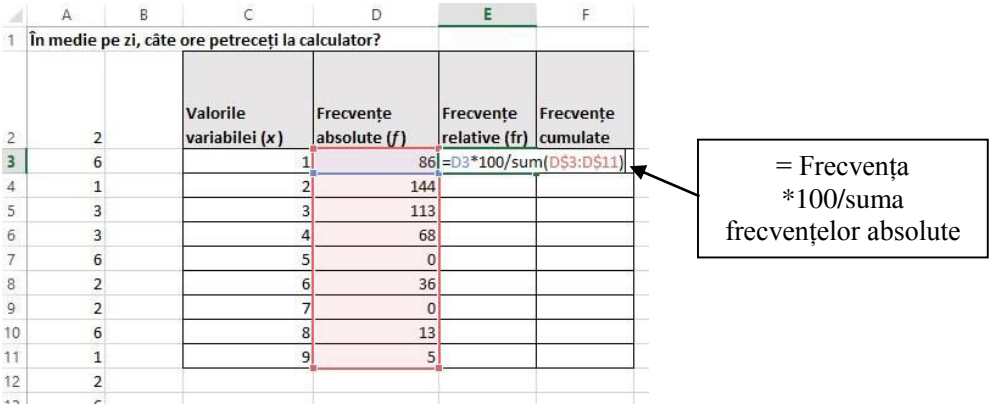
Confirmați apăsând simultan tastele CTRL+SHIFT+ENTER



În medie pe zi, câte ore petreceți la calculator?		Valorile variabilei (x)	Frecvențe absolute (f)
2		1	86
6		2	144
1		3	113
3		4	68
3		5	0
6		6	36
2		7	0
2		8	13
6		9	5
1			
2			
6			
4			

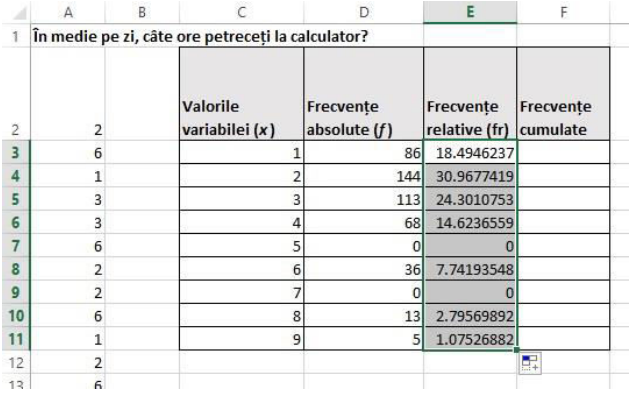
Observăm că două dintre valori nu apar în seria de date: valoarea 5 și valoarea 7 (acestea nu sunt menționate de către subiecți, astfel frecvența lor este 0). Putem să le ștergem din tabel sau putem să le păstrăm. Ele nu vor influența modul de calcul pentru frecvențele relative, iar în cazul celor cumulate le vom șterge din tabel.

Frecvențele relative și frecvențele cumulate le vom calcula în modul prezentat în exemplul anterior. Pentru frecvențele relative vom utiliza formula $fr_i = \frac{f_i}{\sum f_i} * 100$, iar pentru frecvențele cumulate vom însuma pentru frecvența respectivă, suma frecvențelor anterioare.



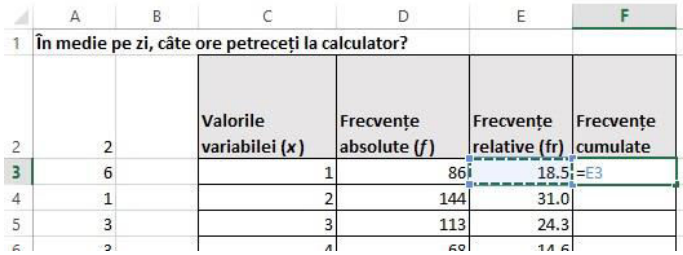
	A	B	C	D	E	F
1	În medie pe zi, câte ore petreceți la calculator?					
2	2		Valorile variabilei (x)	Frecvențe absolute (f)	Frecvențe relative (fr)	Frecvențe cumulate
3	6		1	86	=D3*100/sum(D$3:D$11)	
4	1		2	144		
5	3		3	113		
6	3		4	68		
7	6		5	0		
8	2		6	36		
9	2		7	0		
10	6		8	13		
11	1		9	5		
12	2					

După ce am introdus formula pentru prima frecvență relativă (aferentă primei valori a variabilei), vom utiliza opțiunea auto-umplere. Pentru a putea să multiplicăm formula, este necesar să "blocăm" rândurile aferente referințelor (rândul 3 îl vom bloca prin adăugarea semnului $ înaintea cifrei 3, iar rândul 11 îl vom bloca prin adăugarea semnului $ înaintea cifrei 11).

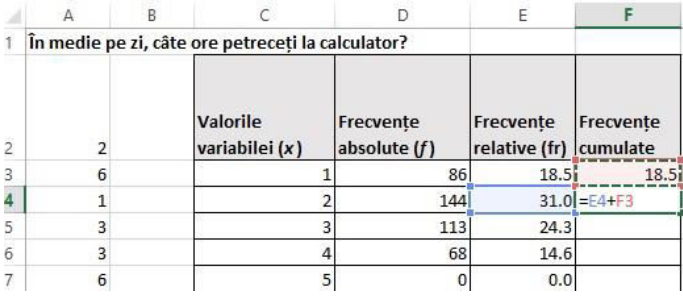


	A	B	C	D	E	F
1	În medie pe zi, câte ore petreceți la calculator?					
2	2		Valorile variabilei (x)	Frecvențe absolute (f)	Frecvențe relative (fr)	Frecvențe cumulate
3	6		1	86	18.4946237	
4	1		2	144	30.9677419	
5	3		3	113	24.3010753	
6	3		4	68	14.6236559	
7	6		5	0	0	
8	2		6	36	7.74193548	
9	2		7	0	0	
10	6		8	13	2.79569892	
11	1		9	5	1.07526882	
12	2					

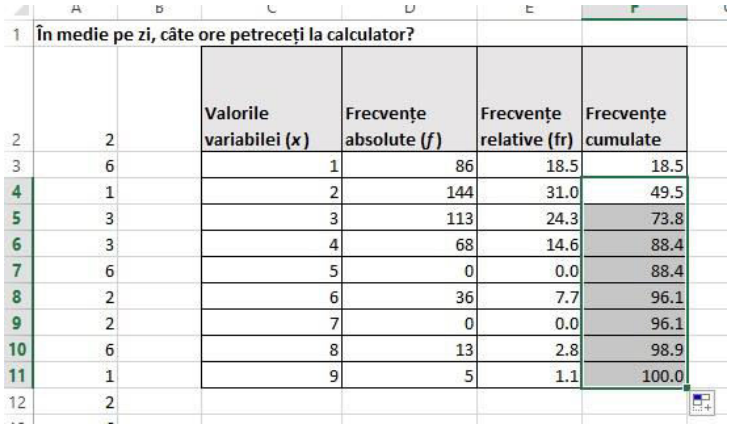
Pentru prima frecvență cumulată vom copia informația regăsită în celula E3 deoarece nu avem alte frecvențe anterioare pe care să le adunăm la aceasta. Apoi, pentru următoarele frecvențe cumulate, vom aduna frecvențele anterioare. Vom folosi și de această dată opțiunea auto-umplere.



	A	B	C	D	E	F
1	În medie pe zi, câte ore petreceți la calculator?					
2	2		Valorile variabilei (*x*)	Frecvențe absolute (*f*)	Frecvențe relative (fr)	Frecvențe cumulate
3	6		1	86	18.5	=E3
4	1		2	144	31.0	
5	3		3	113	24.3	



	A	B	C	D	E	F
1	În medie pe zi, câte ore petreceți la calculator?					
2	2		Valorile variabilei (*x*)	Frecvențe absolute (*f*)	Frecvențe relative (fr)	Frecvențe cumulate
3	6		1	86	18.5	18.5
4	1		2	144	31.0	=E4+F3
5	3		3	113	24.3	
6	3		4	68	14.6	
7	6		5	0	0.0	



	A	B	C	D	E	F
1	În medie pe zi, câte ore petreceți la calculator?					
2	2		Valorile variabilei (*x*)	Frecvențe absolute (*f*)	Frecvențe relative (fr)	Frecvențe cumulate
3	6		1	86	18.5	18.5
4	1		2	144	31.0	49.5
5	3		3	113	24.3	73.8
6	3		4	68	14.6	88.4
7	6		5	0	0.0	88.4
8	2		6	36	7.7	96.1
9	2		7	0	0.0	96.1
10	6		8	13	2.8	98.9
11	1		9	5	1.1	100.0
12	2					

Observăm că în cazul valorilor 5 și 7 (cele care au frecvența 0), se repetă frecvența cumulată a valorii anterioare. Rândurile pe care sunt introduse aceste valori trebuie fie ascunse, fie șterse.

Calculul principalilor indicatori statistici

Este important să ținem cont de faptul că utilizarea indicatorilor statistici pentru descrierea variabilelor este dependentă de nivelul de măsurarea a variabilelor pe care dorim să le analizăm. În cazul variabilelor măsurate la nivel nominal putem să calculăm doar modul (din rândul indicatorilor ai tendinței centrale), pentru variabilele ordinale putem să calculăm modul, mediana (dintre indicatorii tendinței centrale) și quartilele și abaterea interquartilă (dintre indicatorii dispersiei). În cazul

variabilelor numerice putem să calculăm toți indicatorii principali utilizați în statistica descriptivă.

Pentru prezentarea modului de calcul a indicatorilor statistici vom utiliza pentru exemplificare o variabilă numerică. În cele de urmează vom prezenta detaliat modul de calcul utilizând formulele și funcțiile, respectiv vom prezenta principalele funcții statistice și opțiunea ***Descriptive statistics*** (din *Add-In Data Analysis*).

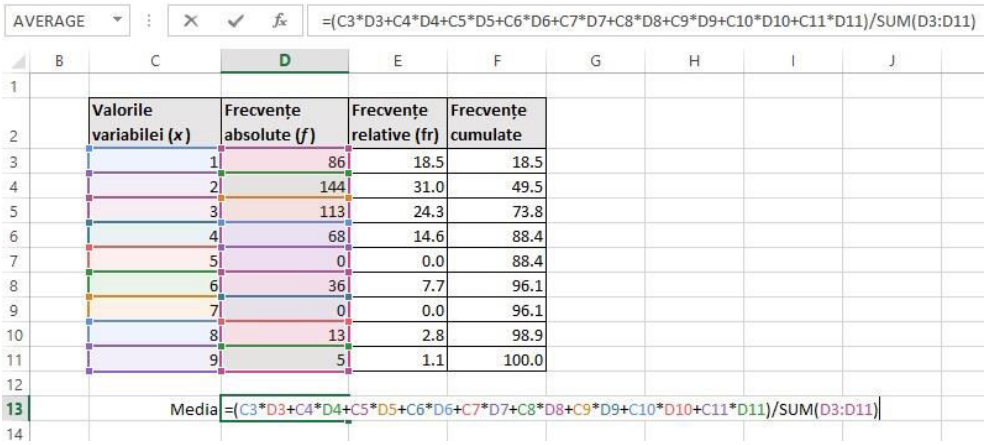
Calcularea principalilor indicatori statistici pe baza tabelului de frecvență

De exemplu: Vom calcula principalii indicatori statistici pornind de la variabila prezentată în exemplul anterior. Variabila este numerică și se referă la numărul de ore petrecut în medie într-o zi de subiecți la calculator. În baza de date sunt înregistrate 465 de cazuri.

Pentru început, vom porni de la tabelul de frecvență realizat în exemplul anterior.

	A	B	C	D	E	F
1	În medie pe zi, câte ore petreceți la calculator?					
2	2		Valorile variabilei (*x*)	Frecvențe absolute (*f*)	Frecvențe relative (fr)	Frecvențe cumulate
3	6		1	86	18.5	18.5
4	1		2	144	31.0	49.5
5	3		3	113	24.3	73.8
6	3		4	68	14.6	88.4
7	6		5	0	0.0	88.4
8	2		6	36	7.7	96.1
9	2		7	0	0.0	96.1
10	6		8	13	2.8	98.9
11	1		9	5	1.1	100.0
12	2					

Vom calcula MEDIA ARITMETICĂ folosind formula $\bar{x} = \frac{\sum x_i * f_i}{\sum f_i}$, unde x_i sunt valorile variabilei, f_i - frecvențele absolute, iar $\Sigma f_i = N$ (numărul total de cazuri; suma frecvențelor).
Pentru a calcula Media putem să introducem direct formula, înmulțind fiecare valoare cu fiecare frecvență și împărțind la suma frecvențelor.



Acest mod de lucru nu este potrivit deoarece alocăm mult timp selecției referințelor și totodată riscăm să greșim în selecția acestora. Pentru a lucra mai rapid, pentru calcularea $\sum x_i * f_i$ (sumei produselor dintre toate valorile și frecvențele) vom utiliza funcția = **SUMPRODUCT**.

Funcția

=SUMPRODUCT(array1, [array2], [array3]…)

Descriere

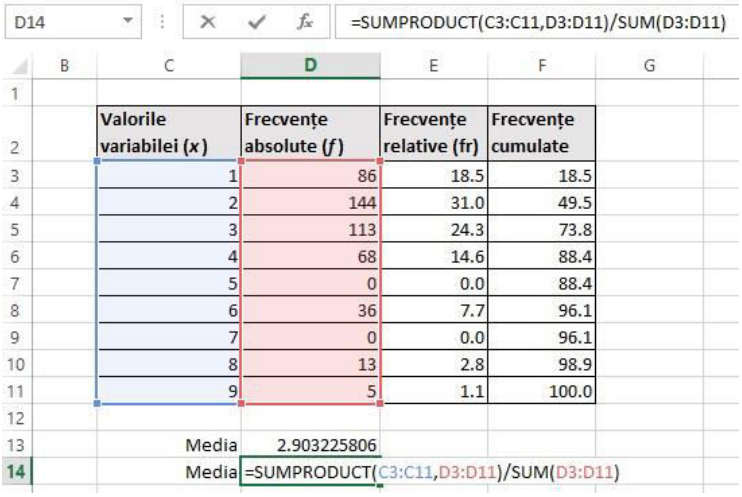
Funcția este folosită pentru calcularea sumelor produselor valorilor regăsite într-o zona de celule.

Argumentele

Array1 – prima zona de celule pe care dorim să le înmulțim și apoi să le adunăm

Array2 – a doua zona de celule pe care dorim să le înmulțim și apoi să le adunăm. Putem să adăugăm coloane mai multe din matrice.

În exemplul nostru vom selecta pentru argumentul ***Array1*** referințele care cuprind valorile variabilelor (C3:C11), iar pentru argumentul ***Array2*** vom selecta celulele care conțin frecvențele (D3:D11).



	Valorile variabilei (x)	Frecvențe absolute (f)	Frecvențe relative (fr)	Frecvențe cumulate
1	1	86	18.5	18.5
2	2	144	31.0	49.5
3	3	113	24.3	73.8
4	4	68	14.6	88.4
5	5	0	0.0	88.4
6	6	36	7.7	96.1
7	7	0	0.0	96.1
8	8	13	2.8	98.9
9	9	5	1.1	100.0
10				
11	Media	2.903225806		
12	Media	=SUMPRODUCT(C3:C11,D3:D11)/SUM(D3:D11)		

Media aritmetică: `=SUMPRODUCT(C3:C11,D3:D11)/SUM(D3:D11)`

Pentru a calcula ABATEREA MEDIE ABSOLUTĂ (media aritmetică a abaterilor individuale de la medie) vom porni de la formula AM $= \frac{\sum |x_i - \bar{x}| \cdot f_i}{N}$, unde x_i sunt valorile variabilei, $\bar{x}$ este media, f_i sunt frecvențele și N este numărul total de cazuri (suma frecvențelor).

Putem să observăm că avem nevoie din nou de suma produselor (vom folosi ca atare funcția =SUMPRODUCT), dar nu vom mai avea valorile variabilei, ci diferențele dintre valorile variabilei și media variabilei. Deoarece sunt utilizate valorile absolute în cazul acestor diferențe (ignorarea semnelor negative), vom folosi funcția =ABS pentru convertirea valorilor negative în valori pozitive.

Funcția

=ABS(number1, [number2], …)

Descriere

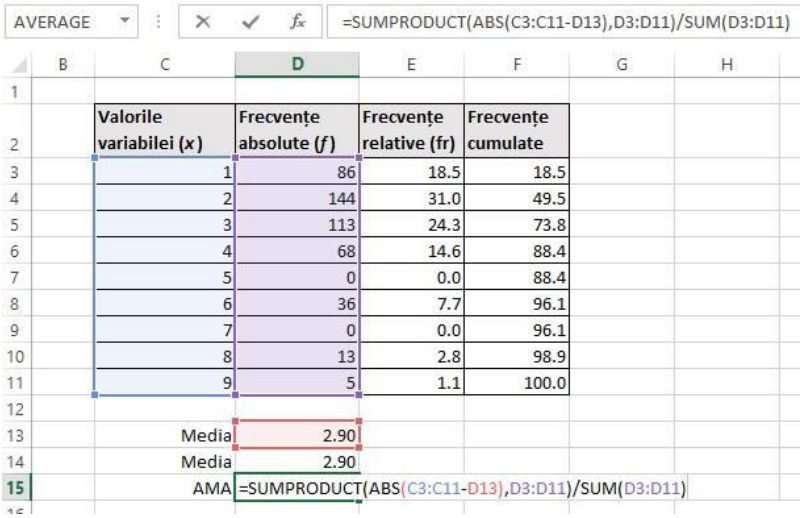
Funcția returnează valoarea absolută a unui număr. Elimină semnul negativ în cazul unui număr.

Argumentele

Number – numărul pe care dorim să îl convertim.

Pe datele noastre, vom avea suma produselor dintre diferențele valorilor variabilei și media acesteia (valorile variabilei se regăsesc în celulele C3:C11, iar media este calculată în celula D13), împărțită la suma frecvențelor (frecvențele se regăsesc în

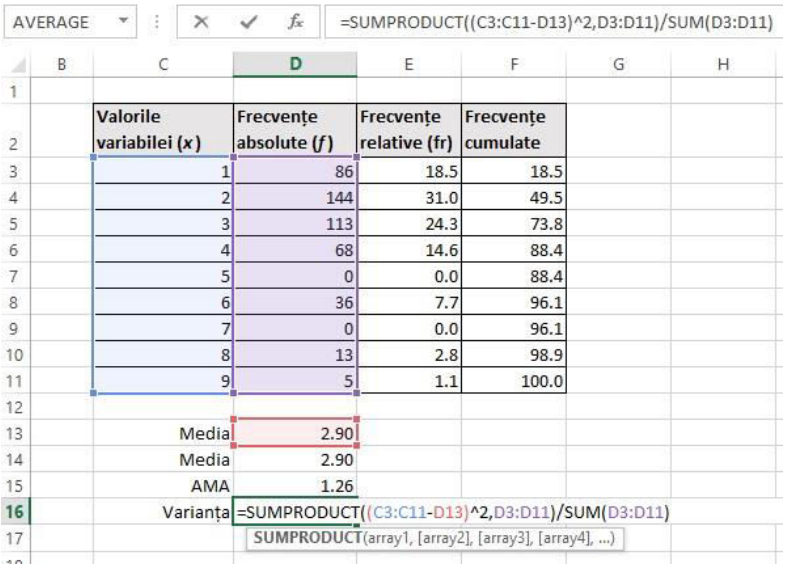
celulele D3:D11). Diferențele sunt convertite în valori absolute, astfel utilizăm funcția =ABS.



Abaterea medie absolută:

```
=SUMPRODUCT(ABS(C3:C11-D13),D3:D11)/SUM(D3:D11)
```

Pentru a calcula VARIANȚA vom avea formula $s^2 = \frac{\sum(x_i - \bar{x})^2 * f_i}{N}$, unde x_i sunt valorile variabilei, $\bar{x}$ este media, f_i sunt frecvențele și N este numărul total de cazuri (suma frecvențelor). Vom ține cont dacă datele sunt obținute la nivel de eșantion (în acest caz din totalul N se va scade valoarea 1). Observăm că diferența față de formula anterioară (a abaterii medii absolute) este reprezentată de ridicarea la pătrat. În această situație nu mai este nevoie să utilizăm funcția =**ABS** pentru convertirea valorilor negative.
Pentru a ridica la pătrat diferențele dintre valorile variabilei și media acesteia vom folosi semnul ^.



Varianța: =SUMPRODUCT((C3:C11-D13)^2,D3:D11)/SUM(D3:D11)

Pentru a calcula ABATEREA STANDARD vom folosi formula $s = \sqrt{s^2} = \sqrt{\frac{\sum(x_i - \bar{x})^2 * f_i}{N-1}}$. Abaterea standard este radical de ordinul doi (rădăcină pătrată) din varianță, astfel, în formula introdusă anterior trebuie să adăugăm radicalul. Pentru aceasta vom folosi funcția =**SQRT**.

Funcția

=SQRT(number1, [number2], …)

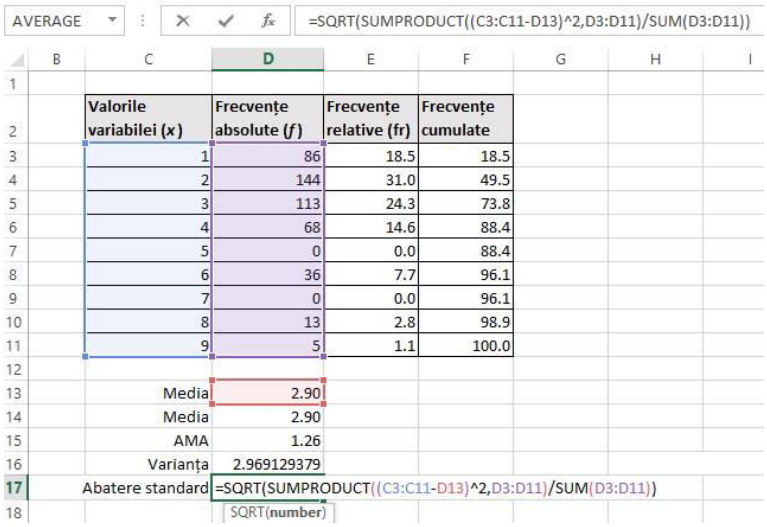
Descriere

Funcția returnează valoarea unei rădăcii pătrate. Prezintă rezultatul extragerii din radical de ordinul 2.

Argumentele

Number – numărul pe care dorim să îl extragem din radical de ordinul 2.

Pentru datele noastre, vom folosi aceeași formula ca în cazul variației,d ar vom adăuga =SQRT.



	Valorile variabilei (x)	Frecvențe absolute (f)	Frecvențe relative (fr)	Frecvențe cumulate
1	86	18.5	18.5
2	144	31.0	49.5
3	113	24.3	73.8
4	68	14.6	88.4
5	0	0.0	88.4
6	36	7.7	96.1
7	0	0.0	96.1
8	13	2.8	98.9
9	5	1.1	100.0

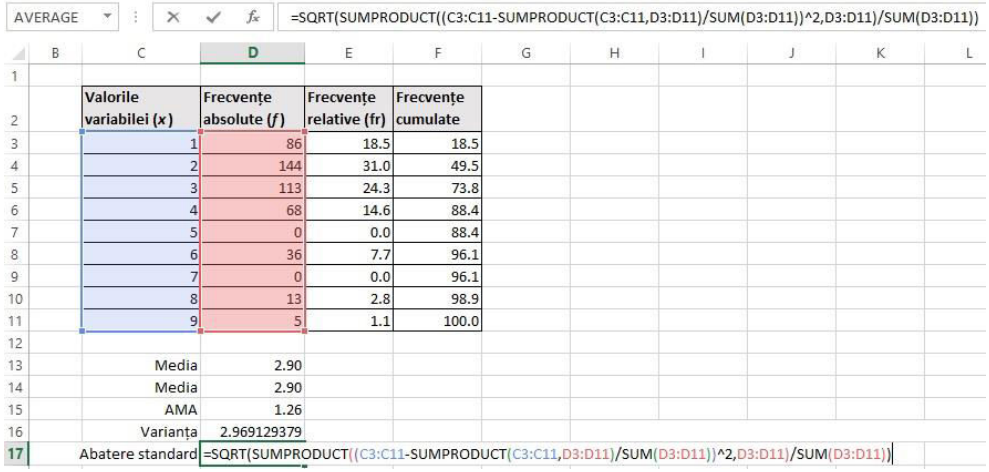
Media	2.90
Media	2.90
AMA	1.26
Varianța	2.969129379
Abatere standard	=SQRT(SUMPRODUCT((C3:C11-D13)^2,D3:D11)/SUM(D3:D11))

SQRT(number)

Abaterea standard:

```
=SQRT(SUMPRODUCT((C3:C11-D13)^2,D3:D11)/SUM(D3:D11))
```

În situația în care nu am calculat anterior într-o altă celulă media aritmetică, putem să scriem direct întreaga formulă pentru abaterea standard. În cazul datelor noastre aceasta este:



	Valorile variabilei (x)	Frecvențe absolute (f)	Frecvențe relative (fr)	Frecvențe cumulate
1	86	18.5	18.5
2	144	31.0	49.5
3	113	24.3	73.8
4	68	14.6	88.4
5	0	0.0	88.4
6	36	7.7	96.1
7	0	0.0	96.1
8	13	2.8	98.9
9	5	1.1	100.0

Media	2.90
Media	2.90
AMA	1.26
Varianța	2.969129379
Abatere standard	=SQRT(SUMPRODUCT((C3:C11-SUMPRODUCT(C3:C11,D3:D11)/SUM(D3:D11))^2,D3:D11)/SUM(D3:D11))

Abaterea standard:

```
=SQRT(SUMPRODUCT((C3:C11-SUMPRODUCT(C3:C11,D3:D11)/SUM(D3:D11))^2,D3:D11)/SUM(D3:D11))
```

Calcularea principalilor indicatori statistici pe baza valorilor variabilei folosind funcțiile statistice

Fără să mai trecem prin etapa realizării tabelului de frecvență, putem să calculăm indicatorii statistici aplicând funcțiile corespunzătoare acestora direct pe valorile variabilei înregistrate în baza de date. Totodată, pentru analiza statistică descriptivă a variabilelor numerice nu este recomandată realizarea tabelului de frecvență, ci este mai important să utilizăm indicatorii statistici care descriu variabila (în special vorbim despre media aritmetică și abaterea standard).

Principalele funcții utilizate pentru calcularea indicatorilor statistici sunt:

Funcția
=AVERAGE(number1, [number2], …)

Descriere

Calculează media aritmetică pe o serie de date. Funcția a fost prezentată în cadrul capitolului privind funcțiile de bază.

Argumentele

Number1…2 – numerele sau referințele pe care dorim să le includem în calcularea mediei aritmetice.

Funcția
=MODE(number1, [number2], …)

Descriere

Calculează modul. Returnează valoarea cea mai frecvent întâlnită într-o serie de date.

Argumentele

Number1…2 – numerele sau referințele din care dorim să identificăm modul.

Funcția
=MEDIAN(number1, [number2], …)

Descriere

Calculează mediana. Returnează valoarea din mijloc dintr-o serie de date. Este acea valoare a unei variabile care împarte seria de date în două părți egale.

Argumentele

Number1…2 – numerele sau referințele pe baza cărora calculăm mediana.

Funcția

=QUARTILE(array, [number2], …)

Descriere

Calculează quartilele. Returnează valoarea aflată la 25%, 50% (mediana) sau 70% dintr-o serie de date.

Argumentele

Array – numerele sau referințele pe baza cărora calculăm quartila; seria de date.

Quart – numărul care indică Quartila: 1 pentru quartila 1 (25%), 2 pentru quartila 2 care este mediana (50%), 3 pentru quartila 3 (75%).

Funcția

=VAR.P(array, [number2], …)

=VAR.S(array, [number2], …)

Descriere

Calculează varianța. **=VAR.P** calculează varianța pentru date obținute la nivelul întregii populații, iar **=VAR.S** calculează varianța pentru date obținute la nivel de eșantion.

Argumentele

Number1…2 – numerele sau referințele din seria de date pe baza cărora dorim să calculăm varianța.

Funcția

=STDEV.P(array, [number2], …)

=STDEV.S(array, [number2], …)

Descriere

Calculează abaterea standard. **=STDEV.P** calculează abaterea standard pentru date obținute la nivelul întregii populații, iar **=STDEV.S** calculează abaterea standard pentru date obținute la nivel de eșantion.

Argumentele

Number1…2 – numerele sau referințele din seria de date pe baza cărora dorim să calculăm abaterea standard.

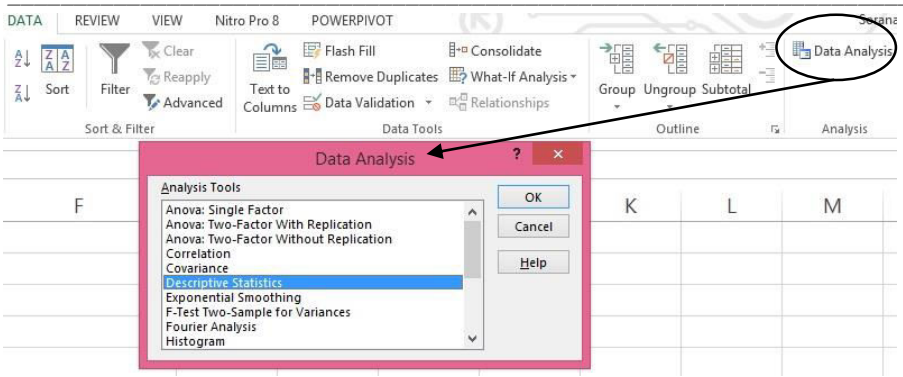
De exemplu: Pornind de la seria de date prezentată în exemplul anterior, vom aplica toate funcțiile pentru calcularea principalilor indicatori. Variabila analizată se referă la numărul de ore petrecut în medie într-o zi de subiecți la calculator (sunt 465 de cazuri în baza de date). Datele sunt introduse pe coloana A. Ținem cont de faptul că datele sunt obținute la nivelul întregii populații, astfel pentru a calcula varianța, respectiv abaterea standard, vom folosi funcțiile **= VAR.P**, respectiv **=STDEV.P**. Funcțiile utilizate pentru a calcula principalii indicatorii statistici pentru a descrie variabila sunt:

	A	B	C	D	E	F
1	**În medie pe zi, câte ore petreceți la calculator?**					
2	2					
3	6					
4	1		Media	2.90		=AVERAGE(A:A)
5	3		Modul	2		=MODE(A:A)
6	3		Mediana	3		=MEDIAN(A:A)
7	6		Quartila 1	2		=QUARTILE(A:A,1)
8	2		Quartila 2	3		=QUARTILE(A:A,2)
9	2		Quartila 3	4		=QUARTILE(A:A,3)
10	6		Varianța	2.97		=VAR.P(A:A)
11	1		Ab.standard	1.72		=STDEV.P(A:A)
12	2					

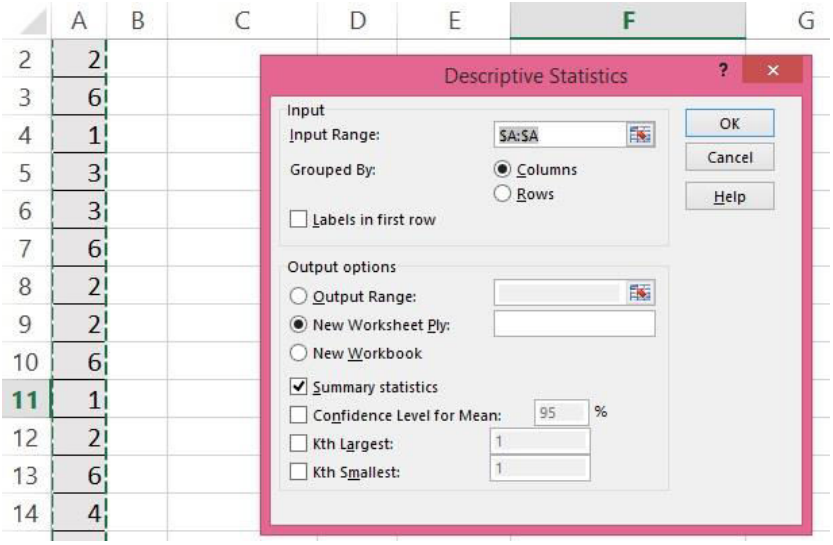
Calcularea principalilor indicatori statistici pe baza valorilor variabilei folosind Add-In Data Analysis

Pentru calcularea principalilor indicatori statistici sau alte operațiuni pentru analiza statistică a datelor poate fi instalată opțiunea ***Data Analysis***. Pentru a accesa acest pachet pentru analiză statistică accesăm meniul FILE > Options. Selectăm Add-Ins > Analysis ToolPak. După instalarea acestui pachet de instrumente, în meniul DATA include *Data Analysis*.

Din fereastră putem selecta *Descriptive statistics* pentru analiza statistică descriptivă pentru o variabilă numerică.



În fereastra aferentă statisticii descriptive (*Descriptive Statistics*) adăugăm la *Input Range* seria de date pe care dorim să o analizăm. Rezultatele (*Output*) putem să alegem să fie adăugate direct în foaia de calcul în care se află seria de date sau într-o altă foaie de calcul sau chiar într-un alt registru de calcul. Selectăm *Summary statistic* pentru a calcula principalii indicatori statistici. Programul este setat pentru pragul de încredere de 95% pentru medie.



Rezultatele (Output) sunt prezentate astfel:

	A	B
1	*Column1*	
2		
3	Mean	2.90323
4	Standard Error	0.07999
5	Median	3
6	Mode	2
7	Standard Deviation	1.72497
8	Sample Variance	2.97553
9	Kurtosis	1.98975
10	Skewness	1.39763
11	Range	8
12	Minimum	1
13	Maximum	9
14	Sum	1350
15	Count	465

Media
Eroarea standard
Mediana
Modul
Abaterea standard
Varianța
Oblicitatea
Boltirea
Amplitudinea
Minimul
Maximul
Suma valorilor
Numărul de cazuri

În utilizarea acestui pachet de instrumente *Data Analysis* trebuie să ținem cont de faptul că indicatorii sunt calculați pornind de la principiul că datele sunt colectate la nivel de eșantion.

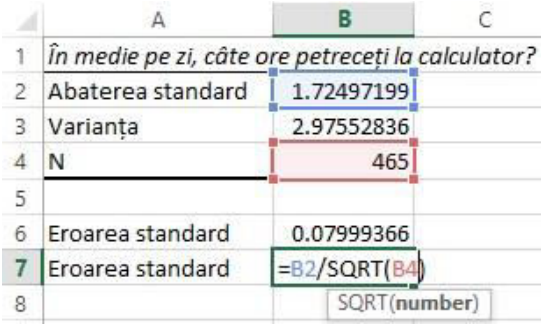
Pornind de la valorile calculate ale Varianței, respectiv Abaterii standard la nivel de eșantion, putem să calculăm EROAREA STANDARD, folosind formula $e = \frac{\sigma}{\sqrt{n}}$ (sau putem să extragem din radical varianța împărțită la suma frecvențelor).

Pentru datele noastre vom utiliza referințele în care am calculat deja abaterea standard, respectiv varianța.

Eroarea standard calculată ca radical din Varianță supra suma frecvențelor:

	A	B	C
1	*În medie pe zi, câte ore petreceți la calculator?*		
2	Abaterea standard	1.72497199	
3	Varianța	2.97552836	
4	N	465	
5			
6	Eroarea standard	=SQRT(B3/B4)	
7	Eroarea standard	SQRT(number)	

Eroarea standard calculată ca abatere standard supra radical din suma frecvențelor:



Pornind de la eroarea standard putem să calculăm intervalul de încredere. În acest interval se regăsește valoarea la nivelul populației:

(valoare eșantion – z*e, valoare eșantion + z*e)

Pentru pragul minim de încredere de 95% (p=0.05), valoarea lui z este 1.96.

Potrivit valorilor obținute pentru intervalul de încredere, putem să afirmăm că sunt 95% șanse ca valoarea la nivelul populației să se încadreze în acest interval.

Putem să calculăm valorile pentru acest interval folosind funcția **=CONFIDENCE.NORM**.

Funcția

=CONFIDENCE.NORM(alpha, standard_dev, size)

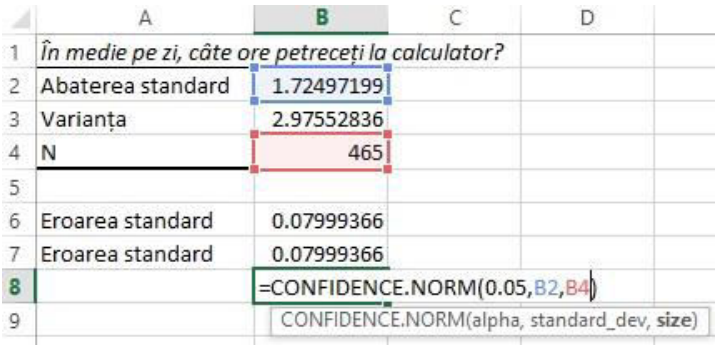
Descriere

Calculează intervalul de încredere pentru o medie pornind de la o distribuție normală.

Argumentele

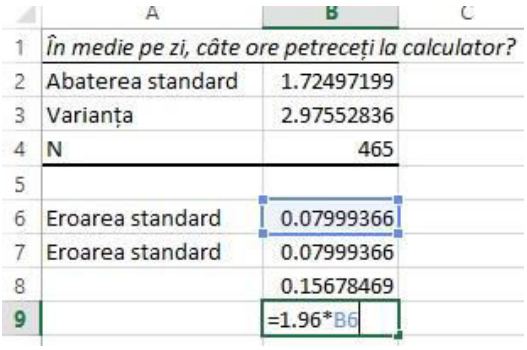
Alpha – pragul de semnificație (p). Pragul minim acceptat în științele sociale este de 0.05 care indică un nivel de încredere de 95%.

Pentru datele din exemplul de mai sus, vom calcula intervalul de încredere pornind de la valoarea returnată folosind funcția **=CONFIDENCE.NORM**.



	A	B	C	D
1	*În medie pe zi, câte ore petreceți la calculator?*			
2	Abaterea standard	1.72497199		
3	Varianța	2.97552836		
4	N	465		
5				
6	Eroarea standard	0.07999366		
7	Eroarea standard	0.07999366		
8		=CONFIDENCE.NORM(0.05,B2,B4)		
9		CONFIDENCE.NORM(alpha, standard_dev, **size**)		

Putem să obținem această valoare și folosind eroarea standard:



	A	B	C
1	*În medie pe zi, câte ore petreceți la calculator?*		
2	Abaterea standard	1.72497199	
3	Varianța	2.97552836	
4	N	465	
5			
6	Eroarea standard	0.07999366	
7	Eroarea standard	0.07999366	
8		0.15678469	
9		=1.96*B6	

Valoarea medie obținută pe aceste date este de 2.90 de ore petrecute la calculator. Astfel, intervalul de încredere va fi:

(2.90-z*e; 2.90+z*e)

(2.90-0.157; 2.90+0.157)

(2.74; 3.03).

Sunt 95% șanse ca valoarea la nivelul populației să se încadreze în acest interval, pornind de la valoarea medie obținută la nivelul eșantionului de 2.90 de ore.

7 ANALIZA STATISTICĂ BIVARIATĂ. TESTAREA RELAȚIEI DE ASOCIERE ÎNTRE VARIABILELE CATEGORIALE

CHISQ.TEST, CHISQ.INV.RT, CHISQ.DIST.RT

De cele mai multe ori prin analiza statistică a datelor dorim să identificăm o relație între variabile. În demersul nostru de a explica diferitele fenomene și procese sociale, nu este suficient să descriem variabilele (să realizăm distribuțiile de frecvențe, de exemplu), ci este nevoie să identificăm relații semnificative din punct de vedere statistic între variabile, încercând astfel să explicăm diferențele între diferite comportamente sau fenomene sociale. Ne interesează în acest sens dacă indivizii statistici tind să se grupeze în anumite categorii ale variabilelor sau mai degrabă distribuția lor este apropiată de una "teoretică".

Testul non-parametric chi-pătrat (*chi-square;* χ^2) este cel mai des utilizat pentru verificarea relației dintre variabilele categoriale. Acesta urmărește identificarea diferențelor semnificative din punct de vedere statistic dintre distribuția observată (empirică; cea obținută în cadrul cercetării) a indivizilor în funcție de două sau mai multe variabile calitative, și distribuția teoretică a acestora în funcție de variabilele analizate.

Realizarea tabelelor de contingență (asociere)

Pentru a testa relația de asociere este nevoie să urmărim distribuția simultană a indivizilor în funcție de două sau mai multe caracteristici calitative. Astfel vom urmări frecvențele obținute pe două sau mai multe variabile nominale sau ordinale. Tabelul care conține aceste frecvențe (observate) se numește tabel de contingență.

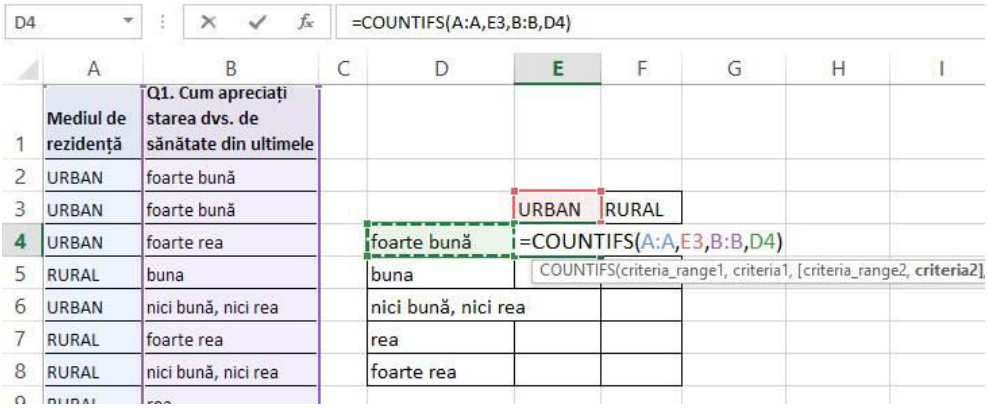
De exemplu: În cadrul unui sondaj cu privire la sănătatea subiectivă realizat pe o populație de 432 de persoane vârstnice, am adresat următoarea întrebare subiecților: *Cum apreciați starea dvs. de sănătate din ultimele 12 luni??*, având următoarele variante de răspuns: " *foarte bună*", "*bună*", " *nici bună, nici rea*", "*rea*", "*foarte*". Alături de această variabilă, am înregistrat datele cu privire la mediul de rezidență: *urban*, *rural*.

Pentru a realiza tabelul de contingență, vom utiliza o structură tip matrice în funcție de categoriile celor două variabile: aprecierea stării de sănătate are 5 categorii, iar

mediul de rezidență are 2 categorii. Tabelul va avea 5 rânduri și 2 coloane; vom avea 10 frecvențe observate (celulele aflate la intersecția rândurilor și coloanelor).

	URBAN	RURAL
foarte bună		
buna		
nici bună, nici rea		
rea		
foarte rea		

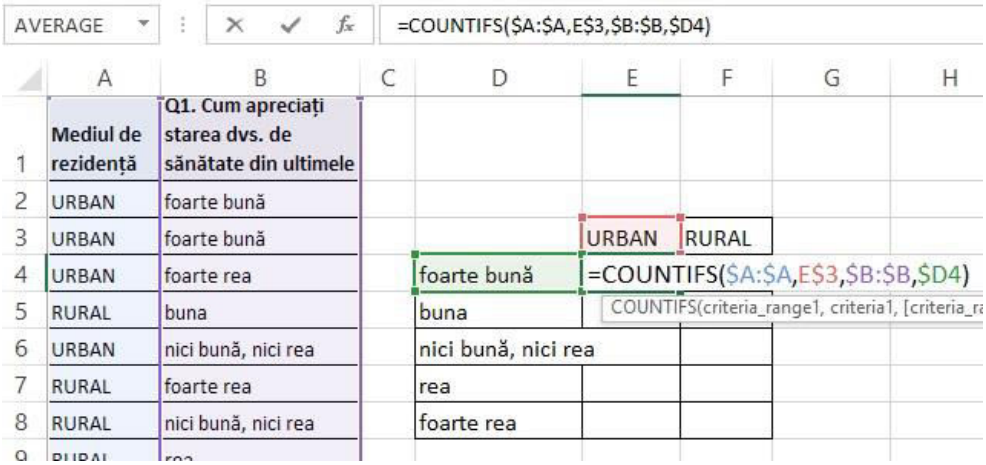
Pentru a identifica frecvențele (observate) vom utiliza funcția **=COUNTIFS** (prezentată în alt capitol), care permite numărarea prin adăugarea mai multor criterii. Criteriile în funcție de care vom face numărarea sunt tocmai categoriile variabilelor. Vom număra: câți subiecți din mediul urban (primul criteriu) au răspuns "foarte bună" (al doilea criteriu) la întrebarea privind aprecierea sănătății. Apoi vom număra câți subiecți din mediul urban au răspuns "bună" la întrebarea privind aprecierea sănătății. Apoi vom număra câți din mediul urban au răspuns " nici bună, nici rea" etc.

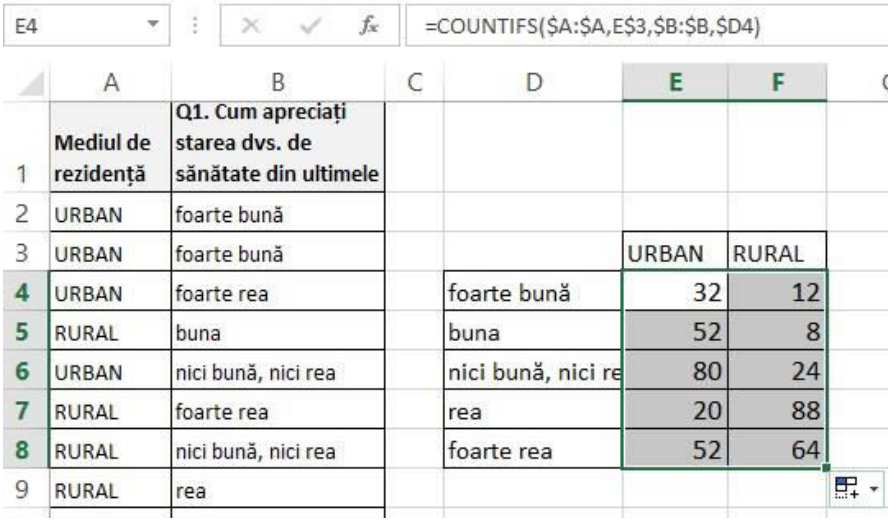


Pentru funcția =**COUNTIFS** vom avea următoarele argumente: pentru *criteria_range1* vom selecta zona celulelor unde se regăsește primul criteriu în funcție de care vom face numărarea, și anume variabila referitoare la mediul de rezidență, apoi la *criteria1* vom selecta pentru început direct din tabel prima categorie a variabilei: URBAN. Apoi vom selecta pentru *criteria_range2* zona celulelor unde se regăsește cel de-al doilea criteriu în funcție de care vom face numărarea, și anume variabila privind starea de sănătate, iar pentru *criteria2* vom selecta prima categorie a variabilei direct din tabel: "foarte bună".

Nu vom introduce funcția de 10 ori (pentru fiecare celulă aferentă frecvențelor observate), deoarece vom folosi opțiunea auto-umplere. Vom multiplica funcția pe rânduri în jos, respectiv pe coloane în dreapta. Pentru a putea realiza acest lucru, este nevoie să blocăm zona celulelor care cuprind datele variabilelor (coloana A, respectiv coloana B), iar în ceea ce privește referințele aferente criteriilor 1 (E3) și 2

(D4) vom bloca: pentru E3 doar rândul 3 deoarece atunci când vom multiplica funcția în jos pe rânduri, nu dorim să treacă referința de pe rândul 3 pe rândul 4, 5 etc., în schimb când vom multiplica funcția pe coloană în dreapta, trebuie să îi permitem referinței să se deplaseze de pe coloana E pe coloana F (unde se regăsește următorul criteriu pentru numărare: RURAL); pentru D4 vom bloca doar coloana D deoarece atunci când vom multiplica funcția în jos pe rânduri trebuie să îi permitem referinței să coboare de pe rândul 4 pe rândul 5, 6 etc. (unde se regăsesc celelalte categorii ale variabilei în funcție de care facem numărarea), în schimb când vom multiplica funcția în dreapta pe coloane nu trebuie să îi permitem referinței să se deplaseze de pe coloana D pe coloana E.





Tabelul de contingență cuprinde frecvențele observate (datele obținute în sondaj pentru cele două variabile analizate).

Testul chi-square (testul chi-pătrat)

Pentru a testa relația de asociere dintre cele două variabile, vom calcula pentru început frecvențele teoretice. Apoi vom compara cele două distribuții: cea observată și cea teoretică.

Pentru fiecare frecvență observată vom avea o frecvență teoretică calculată pe baza frecvențelor marginale pe coloană și a frecvențelor marginale pe rânduri (valorile totale calculate pe coloane, respectiv pe rânduri).

Formula de calcul pentru frecvențele teoretice este:

$$= \frac{frecvența\ marginală\ pe\ coloană * frecvența\ marginală\ pe\ rând}{total}$$

În tabelul care cuprinde frecvențele observate vom calcula valorile totale pe rânduri și pe coloane folosind funcția **=SUM**.

	A	B	C	D	E	F	G
1	**Mediul de rezidență**	**Q1. Cum apreciați starea dvs. de sănătate din ultimele**					
2	URBAN	foarte bună					
3	URBAN	foarte bună			URBAN	RURAL	**TOTAL**
4	URBAN	foarte rea		foarte bună	32	12	**44**
5	RURAL	buna		buna	52	8	**60**
6	URBAN	nici bună, nici rea		nici bună, nici re	80	24	**104**
7	RURAL	foarte rea		rea	20	88	**108**
8	RURAL	nici bună, nici rea		foarte rea	52	64	**116**
9	RURAL	rea		**TOTAL**	=SUM(E4:E8)		**432**
10	URBAN	foarte bună					

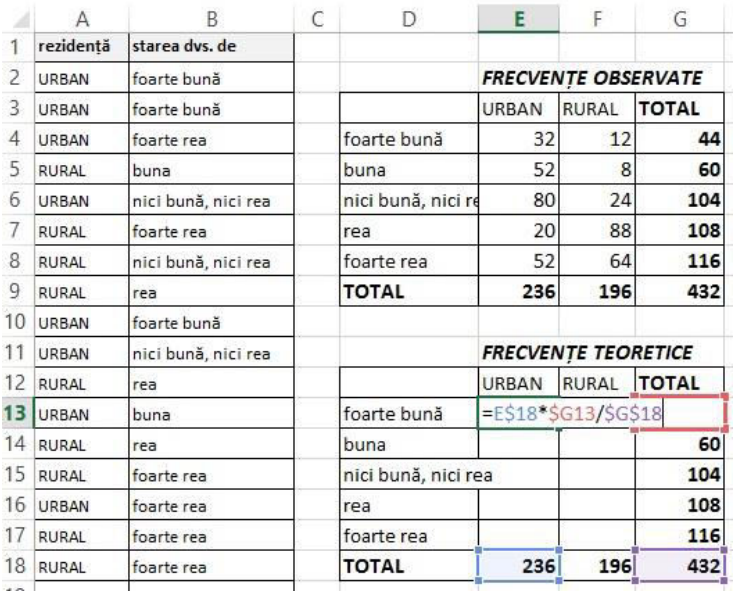
SUM(number1, [number2], ...)

	A	B	C	D	E	F	G
1	**Mediul de rezidență**	**Q1. Cum apreciați starea dvs. de sănătate din ultimele**					
2	URBAN	foarte bună					
3	URBAN	foarte bună			URBAN	RURAL	**TOTAL**
4	URBAN	foarte rea		foarte bună	32	12	=SUM(E4:F4)
5	RURAL	buna		buna	52	8	**60**
6	URBAN	nici bună, nici rea		nici bună, nici re	80	24	**104**
7	RURAL	foarte rea		rea	20	88	**108**
8	RURAL	nici bună, nici rea		foarte rea	52	64	**116**
9	RURAL	rea		**TOTAL**	**236**	**196**	**432**
10	URBAN	foarte bună					

Pornind de la aceste valori marginale vom calcula frecvențele teoretice păstrând structura tabelului.

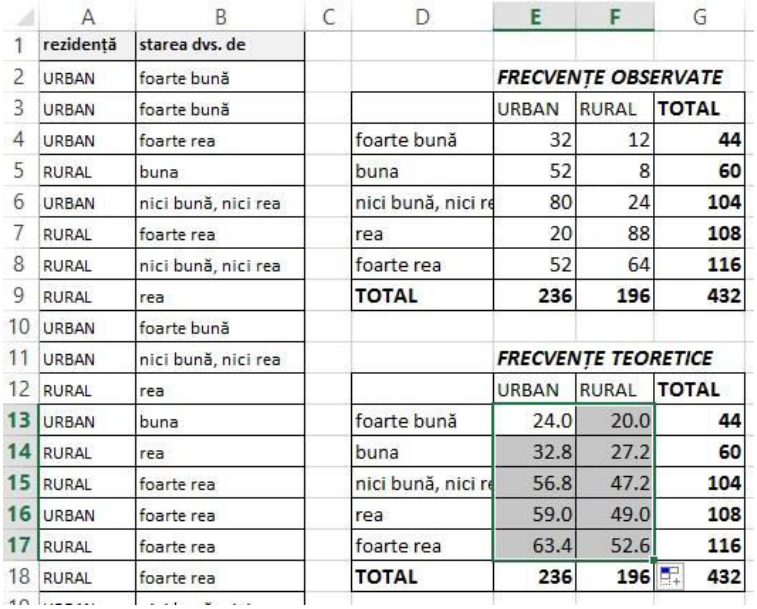
	A	B	C	D	E	F	G
1	**rezidență**	**starea dvs. de**					
2	URBAN	foarte bună			***FRECVENȚE OBSERVATE***		
3	URBAN	foarte bună			URBAN	RURAL	**TOTAL**
4	URBAN	foarte rea		foarte bună	32	12	**44**
5	RURAL	buna		buna	52	8	**60**
6	URBAN	nici bună, nici rea		nici bună, nici re	80	24	**104**
7	RURAL	foarte rea		rea	20	88	**108**
8	RURAL	nici bună, nici rea		foarte rea	52	64	**116**
9	RURAL	rea		**TOTAL**	**236**	**196**	**432**
10	URBAN	foarte bună					
11	URBAN	nici bună, nici rea			***FRECVENȚE TEORETICE***		
12	RURAL	rea			URBAN	RURAL	**TOTAL**
13	URBAN	buna		foarte bună	=E18*G13/G18		**44**
14	RURAL	rea		buna			**60**
15	RURAL	foarte rea		nici bună, nici rea			**104**
16	URBAN	foarte rea		rea			**108**
17	RURAL	foarte rea		foarte rea			**116**
18	RURAL	foarte rea		**TOTAL**	**236**	**196**	**432**

Vom relativiza referințele pentru a putea folosi opțiunea auto-umplere. În cazul referinței E18 vom bloca rândul 18 (prin multiplicarea formulei în dreapta permitem referinței să se deplaseze pe coloana F, dar când multiplicăm pe rânduri blocăm rândul 18 ca referința să nu coboare pe rândul 19 etc.), în cazul referinței G13 vom bloca coloana G (permitem referinței să coboare prin multiplicare de pe rândul 13 pe rândurile următoare, dar blocăm coloana ca atunci când multiplicăm spre dreapta referința să nu se deplaseze pe coloana H), iar pentru referința G18 blocăm și coloana G și rândul 18 (întotdeauna vom împărți la numărul total).



	A	B	C	D	E	F	G
1	**rezidență**	**starea dvs. de**					
2	URBAN	foarte bună			***FRECVENȚE OBSERVATE***		
3	URBAN	foarte bună			URBAN	RURAL	**TOTAL**
4	URBAN	foarte rea		foarte bună	32	12	**44**
5	RURAL	buna		buna	52	8	**60**
6	URBAN	nici bună, nici rea		nici bună, nici re	80	24	**104**
7	RURAL	foarte rea		rea	20	88	**108**
8	RURAL	nici bună, nici rea		foarte rea	52	64	**116**
9	RURAL	rea		**TOTAL**	**236**	**196**	**432**
10	URBAN	foarte bună					
11	URBAN	nici bună, nici rea			***FRECVENȚE TEORETICE***		
12	RURAL	rea			URBAN	RURAL	**TOTAL**
13	URBAN	buna		foarte bună	=E$18*$G13/$G$18		
14	RURAL	rea		buna			**60**
15	RURAL	foarte rea		nici bună, nici rea			**104**
16	URBAN	foarte rea		rea			**108**
17	RURAL	foarte rea		foarte rea			**116**
18	RURAL	foarte rea		**TOTAL**	**236**	**196**	**432**

Tabelul rezultat prin multiplicarea formulei cuprinde frecvențele teoretice.



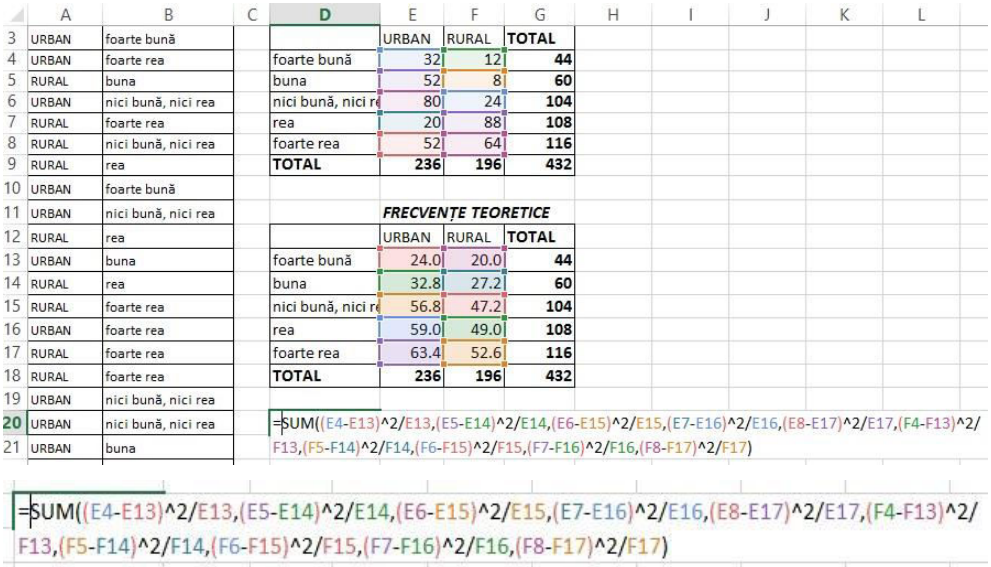
	A	B	C	D	E	F	G
1	**rezidență**	**starea dvs. de**					
2	URBAN	foarte bună			***FRECVENȚE OBSERVATE***		
3	URBAN	foarte bună			URBAN	RURAL	**TOTAL**
4	URBAN	foarte rea		foarte bună	32	12	**44**
5	RURAL	buna		buna	52	8	**60**
6	URBAN	nici bună, nici rea		nici bună, nici re	80	24	**104**
7	RURAL	foarte rea		rea	20	88	**108**
8	RURAL	nici bună, nici rea		foarte rea	52	64	**116**
9	RURAL	rea		**TOTAL**	**236**	**196**	**432**
10	URBAN	foarte bună					
11	URBAN	nici bună, nici rea			***FRECVENȚE TEORETICE***		
12	RURAL	rea			URBAN	RURAL	**TOTAL**
13	URBAN	buna		foarte bună	24.0	20.0	**44**
14	RURAL	rea		buna	32.8	27.2	**60**
15	RURAL	foarte rea		nici bună, nici re	56.8	47.2	**104**
16	URBAN	foarte rea		rea	59.0	49.0	**108**
17	RURAL	foarte rea		foarte rea	63.4	52.6	**116**
18	RURAL	foarte rea		**TOTAL**	**236**	**196**	**432**

Pe cele două distribuții vom aplica testul chi-pătrat.

Pentru a calcula valoarea lui chi-pătrat vom utiliza formula

$$\chi^2=\sum_i^n \frac{(O_i-T_i)2}{T_i}$$ unde O – frecvențele observate, T – frecvențele teoretice.

Putem să scriem formula direct într-o celulă folosind referințele corespunzătoare din tabelele care cuprind cele două distribuții:



	A	B	C	D	E	F	G	H	I	J	K	L
3	URBAN	foarte bună			URBAN	RURAL	TOTAL					
4	URBAN	foarte rea		foarte bună	32	12	44					
5	RURAL	buna		buna	52	8	60					
6	URBAN	nici bună, nici rea		nici bună, nici r	80	24	104					
7	RURAL	foarte rea		rea	20	88	108					
8	RURAL	nici bună, nici rea		foarte rea	52	64	116					
9	RURAL	rea		TOTAL	236	196	432					
10	URBAN	foarte bună										
11	URBAN	nici bună, nici rea			FRECVENȚE TEORETICE							
12	RURAL	rea			URBAN	RURAL	TOTAL					
13	URBAN	buna		foarte bună	24.0	20.0	44					
14	RURAL	rea		buna	32.8	27.2	60					
15	RURAL	foarte rea		nici bună, nici r	56.8	47.2	104					
16	URBAN	foarte rea		rea	59.0	49.0	108					
17	RURAL	foarte rea		foarte rea	63.4	52.6	116					
18	RURAL	foarte rea		TOTAL	236	196	432					
19	URBAN	nici bună, nici rea										
20	URBAN	nici bună, nici rea		=SUM((E4-E13)^2/E13,(E5-E14)^2/E14,(E6-E15)^2/E15,(E7-E16)^2/E16,(E8-E17)^2/E17,(F4-F13)^2/								
21	URBAN	buna		F13,(F5-F14)^2/F14,(F6-F15)^2/F15,(F7-F16)^2/F16,(F8-F17)^2/F17)								

```
=SUM((E4-E13)^2/E13,(E5-E14)^2/E14,(E6-E15)^2/E15,(E7-E16)^2/E16,(E8-E17)^2/E17,(F4-F13)^2/
F13,(F5-F14)^2/F14,(F6-F15)^2/F15,(F7-F16)^2/F16,(F8-F17)^2/F17)
```

Rezultatul obținut este 112.83. Aceasta este valoarea lui chi-pătrat calculat. Pentru a verifica relația de semnificație statistică, trebuie să comparăm această valoare cu cea regăsită în tabelul statistic care cuprinde valorile lui chi-pătrat tabelar.

Pentru a putea să identificăm valoarea lui chi-pătrat tabelar, avem nevoie de gradele de libertate. Acestea se calculează folosind formula:

$$= (număr\ de\ coloane - 1) * (număr\ de\ rânduri - 1)$$

Vom folosi funcțiile **=COLUMNS**, respectiv **=ROWS** pentru a calcula gradele de libertate.

Funcția

=COLUMNS(array)

Descriere

Funcția este folosită pentru a calcula numărul de coloane dintr-o serie de date.

Argumentele

Array – zona de celule care cuprinde seria de date.

Funcția

= ROWS(array)

Descriere

Funcția este folosită pentru a calcula numărul de rânduri dintr-o serie de date.

Argumentele

Array – zona de celule care cuprinde seria de date.

Pentru exemplul nostru, formula pentru calcularea gradelor de libertate va fi:

	C	D	E	F	G
11			***FRECVENȚE TEORETICE***		
12			URBAN	RURAL	**TOTAL**
13		foarte bună	24.0	20.0	**44**
14		buna	32.8	27.2	**60**
15		nici bună, nici re	56.8	47.2	**104**
16		rea	59.0	49.0	**108**
17		foarte rea	63.4	52.6	**116**
18		**TOTAL**	**236**	**196**	**432**
19					
20					
21		=(COLUMNS(E13:F17)-1)*(ROWS(E13:F17)-1)			

Rezultatul este 4 grade de libertate. Este suficient să căutăm în tabelul statistic pentru 4 grade de libertate, valoarea lui chi-pătrat pentru probabilitatea p=0.05 pentru a identifica dacă în cazul celor două variabile analizate există o relație de asociere. Dacă chi-pătrat$_{calculat}$ > chi-pătrat$_{tabelar/critic}$ atunci diferența este semnificativă din punct de vedere statistic. Astfel, între cele două variabile există o relație de asociere. Dacă chi-pătrat$_{calculat}$ < chi-pătrat$_{tabelar/critic}$ atunci diferența nu este semnificativă statistic.

În tabelul statistic valoarea lui chi-pătrat$_{tabelar/critic}$ este 9.48773 pentru 4 grade de libertate.

g.l./aria	0.995	0.99	0.975	0.95	0.9	0.75	0.5	0.25	0.1	0.05	0.025	0.01	0.005
1	0.00004	0.00016	0.00098	0.00393	0.01579	0.10153	0.45494	1.3233	2.70554	3.84146	5.02389	6.6349	7.87944
2	0.01003	0.0201	0.05064	0.10259	0.21072	0.57536	1.38629	2.77259	4.60517	5.99146	7.37776	9.21034	10.5966
3	0.07172	0.11483	0.2158	0.35185	0.58437	1.21253	2.36597	4.10834	6.25139	7.81473	9.3484	11.34487	12.8381
4	0.20699	0.29711	0.48442	0.71072	1.06362	1.92256	3.35669	5.38527	7.77944	9.48773	11.14329	13.2767	14.8602
5	0.41174	0.5543	0.83121	1.14548	1.61031	2.6746	4.35146	6.62568	9.23636	11.0705	12.8325	15.08627	16.7496

Pentru datele noastre 112.83>9.487. Astfel, chi-pătrat$_{calculat}$ > chi-pătrat$_{tabelar/critic,}$ ca atare diferența este semnificativă din punct de vedere statistic. Există o relație de asociere între cele două variabile: starea de sănătate subiectivă și mediul de rezidență.

Pornind de la distribuția frecvențelor observate și a celor teoretice, putem să aplicăm direct funcția pentru testul chi-square pentru a verifica relația de semnificație statistică.

Funcția

=CHISQ.TEST(actual_range, expected_range)

Descriere

Funcția este folosită pentru a testa relația de semnificație (de independență) dintre două distribuții, una observată și una teoretică/așteptată. Funcția returnează valoarea probabilității (valoarea lui p) ca cele două distribuții să fie independente, și anume distribuția observată să nu difere în mod semnificativ de cea teoretică. Valoarea lui p ne spune care este probabilitatea să obținem o anumită valoare pentru chi-pătrat în cazul în care ipoteza nulă este adevărată, și anume cele două distribuții sunt independente. Dacă valoarea obținută este mai mică decât pragul minim de 0.05 atunci această ipoteză de nul, care presupune independența distribuțiilor, este respinsă.

Argumentele

Actual_range – zona celulelor care cuprind frecvențele observate

Expected_range – zona celulelor care cuprind frecvențele teoretice.

De exemplu: Pe datele utilizate în exemplul anterior, vom folosi funcția **=CHISQ.TEST**. Pentru argumentele funcției vom selecta distribuția frecvențelor observate, respectiv distribuția frecvențelor teoretice.

	C	D	E	F	G	H	I	J	K
1									
2			***FRECVENȚE OBSERVATE***						
3			URBAN	RURAL	**TOTAL**		4	gl	
4		foarte bună	32	12	**44**				
5		buna	52	8	**60**		=CHISQ.TEST(E4:F8,E13:F17)		
6		nici bună, nici re	80	24	**104**				
7		rea	20	88	**108**				
8		foarte rea	52	64	**116**				
9		**TOTAL**	**236**	**196**	**432**				
10									
11			***FRECVENȚE TEORETICE***						
12			URBAN	RURAL	**TOTAL**				
13		foarte bună	24.0	20.0	**44**				
14		buna	32.8	27.2	**60**				
15		nici bună, nici re	56.8	47.2	**104**				
16		rea	59.0	49.0	**108**				
17		foarte rea	63.4	52.6	**116**				
18		**TOTAL**	**236**	**196**	**432**				

Rezultatul obținut este 0.00 (mai exact 0.000000000000000000000018). Aceasta este valoarea probabilității ca cele două distribuții să fie independente. Comparăm această valoarea cu pragul minim statistic p=0.05 (pentru încredere de 95%). Valoarea obținută prin aplicarea funcției **=CHISQ.TEST** este mai mică decât pragul minim de 0.05 (0.00<0.05), astfel putem concluziona că între cele două variabile analizate există o relație semnificativă din punct de vedere statistic. Există o asociere între starea de sănătate subiectivă și mediul de rezidență.

Pentru a afla valoarea lui chi-pătrat calculat pentru această probabilitate vom utiliza funcția **=CHISQ.INV.RT**.

Funcția

=CHISQ.INV.RT(probability, deg_freedom)

Descriere

Funcția returnează valoarea inversă la dreapta a probabilității distribuției chi-pătrat.

Argumentele

Probability – probabilitatea distribuției

Deg_freedon – gradele de libertate.

De exemplu: Pentru rezultatele obținute în exemplul anterior, vom utiliza funcția **= CHISQ.INV.RT**. Vom adăuga pentru probabilitate valoarea obținută anterior aplicând funcția **=CHISQ.TEST**, care returnează valoarea lui p pentru compararea independenței celor două distribuții (cea observată și cea teoretică). Pentru gradele de libertate, fie vom scrie valoarea 4, fie vom selecta referința în care am calculat gradele de libertate aplicând formula =(**COLUMNS**(…) -1)*(**ROWS**(…)-1).

	D	E	F	G
2		***FRECVENȚE OBSERVATE***		
3		URBAN	RURAL	**TOTAL**
4	foarte bună	32	12	**44**
5	buna	52	8	**60**
6	nici bună, nici re	80	24	**104**
7	rea	20	88	**108**
8	foarte rea	52	64	**116**
9	**TOTAL**	**236**	**196**	**432**

I3: 4 gl

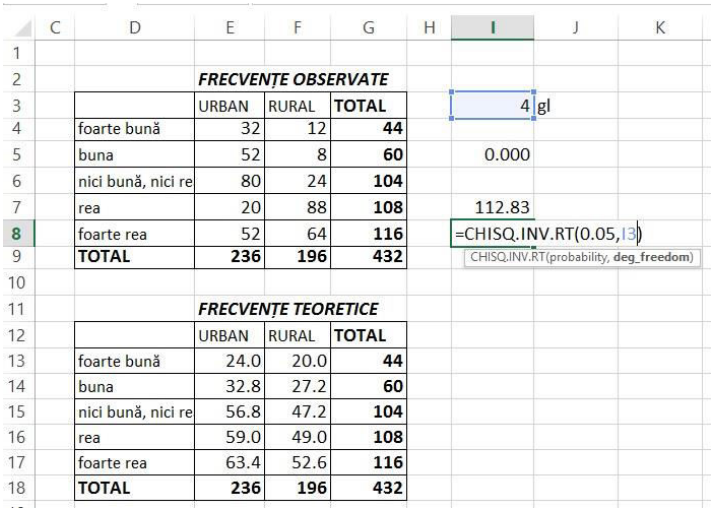
I5: 0.000

I7: =CHISQ.INV.RT(I5,I3)

CHISQ.INV.RT(probability, **deg_freedom**)

Vom obține valoarea 112.83 care este tocmai valoarea lui chi-pătrat$_{calculat}$ pentru datele din cele două distribuții.

Dacă dorim să obținem valoarea lui chi-pătrat$_{tabelar}$, pentru a putea să o comparăm cu valoarea lui cu chi-pătrat$_{calculat}$, vom aplica din nou funcția **=CHISQ.INV.RT**, iar pentru probabilitate vom trece valoarea minimă pentru p de 0.05.



	FRECVENȚE OBSERVATE			
	URBAN	RURAL	**TOTAL**	
foarte bună	32	12	**44**	
buna	52	8	**60**	
nici bună, nici re	80	24	**104**	
rea	20	88	**108**	
foarte rea	52	64	**116**	
TOTAL	**236**	**196**	**432**	
	FRECVENȚE TEORETICE			
	URBAN	RURAL	**TOTAL**	
foarte bună	24.0	20.0	**44**	
buna	32.8	27.2	**60**	
nici bună, nici re	56.8	47.2	**104**	
rea	59.0	49.0	**108**	
foarte rea	63.4	52.6	**116**	
TOTAL	**236**	**196**	**432**	

Rezultatul obținut (9.488) este valoarea lui chi-pătrat tabelar pentru probabilitatea p=0.05 și 4 grade de libertate. Comparăm această valoare cu cea a lui chi-pătrat$_{calculat}$ (9.488<112.83) și vom observa că există o relație semnificativă din punct de vedere statistic, există o asociere între cele două variabile.

Aplicând funcția =**CHISQ.DIST.RT** pe cele două valori ale lui chi-pătrat (calculat și tabelar) obținem tocmai valorile lui p aferente lor (0.00, respectiv 0.05).

Funcția

= CHISQ.DIST.RT(x, deg_freedom)

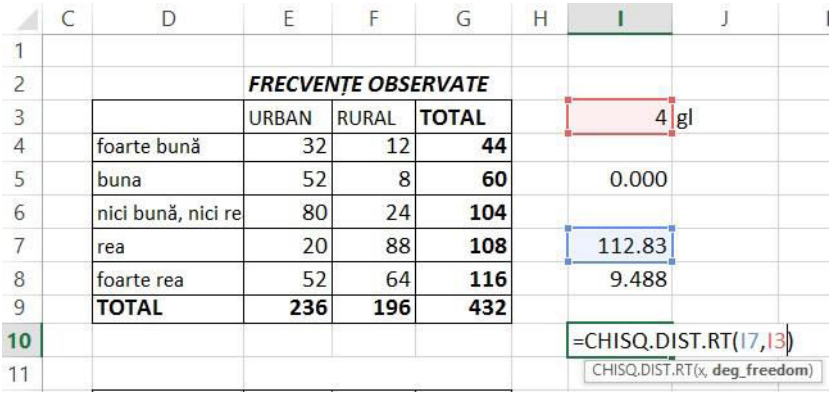
Descriere

Funcția returnează probabilitatea inversă la dreapta a distribuției chi-pătrat.

Argumentele

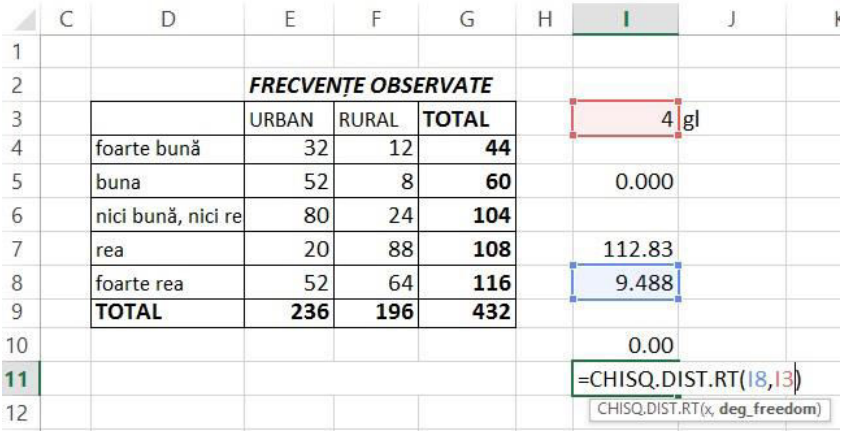
x – valoarea pentru care calculăm probabilitatea

Deg_freedon – gradele de libertate.



	D	E	F	G
2		***FRECVENȚE OBSERVATE***		
3		URBAN	RURAL	**TOTAL**
4	foarte bună	32	12	**44**
5	buna	52	8	**60**
6	nici bună, nici re	80	24	**104**
7	rea	20	88	**108**
8	foarte rea	52	64	**116**
9	**TOTAL**	**236**	**196**	**432**

I3: 4 gl
I5: 0.000
I7: 112.83
I8: 9.488
I10: =CHISQ.DIST.RT(I7,I3)
CHISQ.DIST.RT(x, **deg_freedom**)



	D	E	F	G
2		***FRECVENȚE OBSERVATE***		
3		URBAN	RURAL	**TOTAL**
4	foarte bună	32	12	**44**
5	buna	52	8	**60**
6	nici bună, nici re	80	24	**104**
7	rea	20	88	**108**
8	foarte rea	52	64	**116**
9	**TOTAL**	**236**	**196**	**432**

I3: 4 gl
I5: 0.000
I7: 112.83
I8: 9.488
I10: 0.00
I11: =CHISQ.DIST.RT(I8,I3)
CHISQ.DIST.RT(x, **deg_freedom**)

8 REPREZENTAREA GRAFICĂ A DATELOR

Sunt multe situațiile în care avem nevoie de timp doar pentru a "citi" datele inițiale, ca ulterior să decidem cu privire la includerea acestora în diferite tipuri de analize. Această situație se datorează în special cantității mari de date cuprinsă într-o bază de date. În astfel de cazuri, decizia de a prezenta datele sub formă grafică este una potrivită. Prin includerea datelor într-o singură reprezentare grafică se scurtează timpul de citire a informațiilor. Este nevoie de o singură privire asupra unui grafic pentru a înțelege cum se prezintă datele, ce ne indică ele. Totul, evident, cu mențiunea ca graficul să fie unul realizat corect și să conțină toate datele relevante.

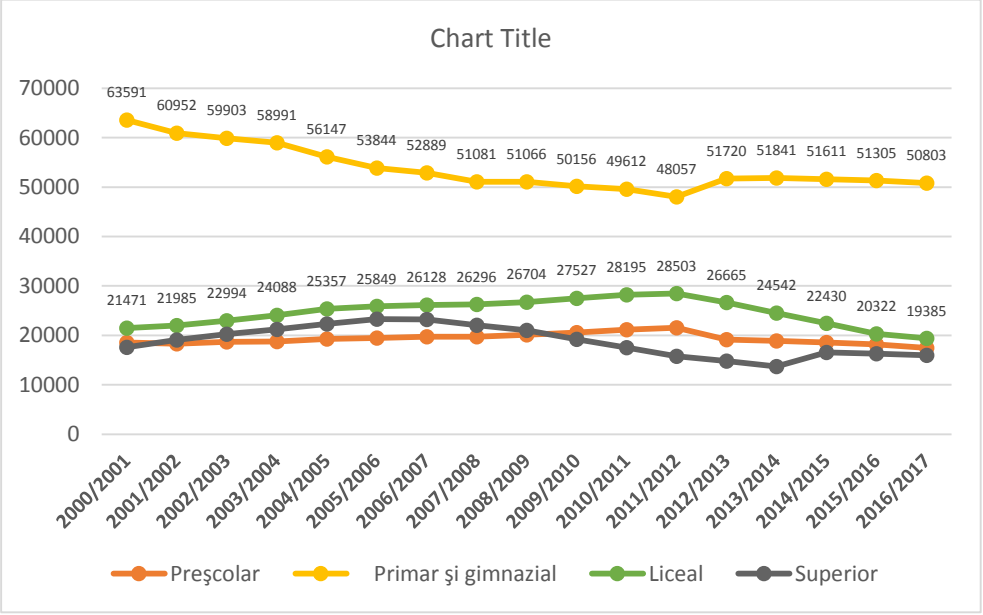
De exemplu: Este destul de dificil să citim datele prezentate în tabelul de mai jos, dintr-o singură privire, și să înțelegem ce ne transmit ele. Datele prezintă evoluția efectivelor școlare în perioada 2000-2016, în județul Bihor. Fiind vorba de schimbări ale cifrelor înregistrate în fiecare an școlar, este important pentru noi să identificăm care este trendul – este o evoluție ascendentă sau descendentă?.

Tabel: *Populaţia şcolară, pe niveluri de educaţie*

Judeţul Bihor	**Preşcolar**	**Primar şi gimnazial**	**Liceal**	**Superior**
2000/2001	18558	63591	21471	17557
2001/2002	18325	60952	21985	19090
2002/2003	18688	59903	22994	20267
2003/2004	18751	58991	24088	21197
2004/2005	19258	56147	25357	22304
2005/2006	19481	53844	25849	23261
2006/2007	19732	52889	26128	23230
2007/2008	19752	51081	26296	22075
2008/2009	20105	51066	26704	21013
2009/2010	20601	50156	27527	19182
2010/2011	21184	49612	28195	17497
2011/2012	21556	48057	28503	15790
2012/2013	19162	51720	26665	14795
2013/2014	18891	51841	24542	13710
2014/2015	18537	51611	22430	16540
2015/2016	18187	51305	20322	16304
2016/2017	17472	50803	19385	15949

(Sursa datelor Direcţia Judeţeană de Statistică BIHOR, www.bihor.insse.ro)

Dacă toate datele le adăugăm într-un grafic, avem într-o singură imagine liniile care ne indică evoluţia numărului de persoane înregistrate pe toate cele 4 niveluri de studii. Această reprezentare grafică este sugestivă, datele sunt uşor de citit şi uşor de interpretat.



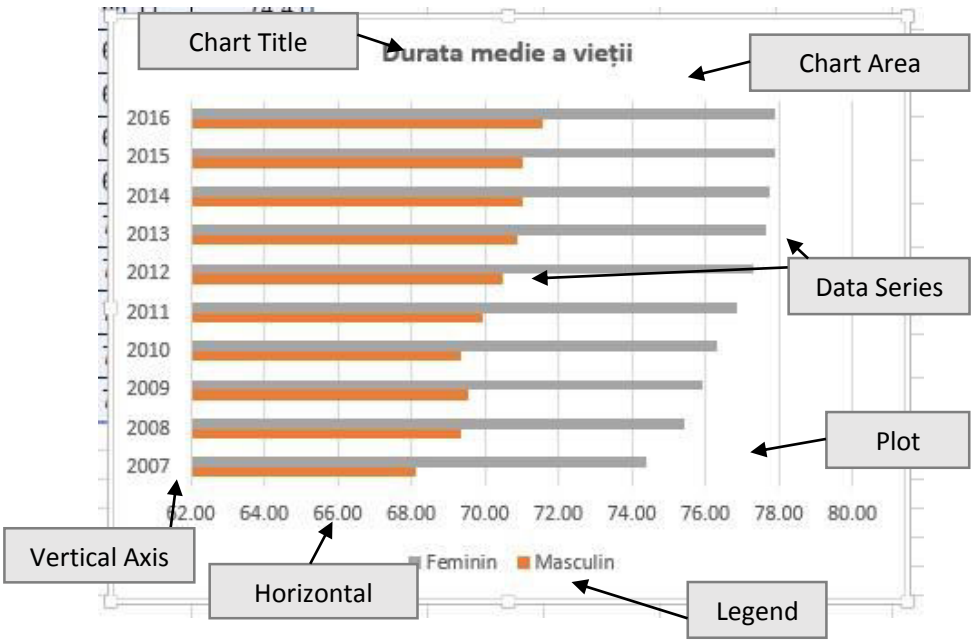
Grafic: *Populaţia şcolară, pe niveluri de educaţie*

(Sursa datelor Direcţia Judeţeană de Statistică BIHOR, www.bihor.insse.ro)

Care sunt elementele principale ale unui grafic?

Atunci când realizăm un grafic folosind programul Excel trebuie să ne asigurăm că toatele elementele sunt corect prezentate. Modul în care sunt formatate aceste elemente permite "citirea" cu uşurinţă a informaţiilor incluse în grafic.

Elementele standard ale unui grafic sunt: suprafaţa graficului (*Chart Area*), seria datelor – valorile categoriilor incluse în grafic (*Chart Data Series*), axele (*Axis*) – în cazul graficelor pe două dimeniuni este axa pe orizontală şi cea pe verticală, suprafaţa marcată de axe (*Plot Area*) şi legenda (*Legend*).



Cum realizăm un grafic în Excel?

Este extrem de simplu să realizăm grafice folosind programul Excel. Însă, după cum am menționat anterior, cel mai important este să ne asigurăm că graficul este realizat corect și cuprinde toate informațiile necesare pentru a interpreta datele.

Pentru început accesăm meniul **INSERT>Charts**.

Realizarea unui grafic pornește de la selecția datelor pe care dorim să le prezentăm sub această formă. Nu vom selecta doar coloana/coloanele sau rândul/rândurile care conțin numerele, ci vom adăuga în selecție și celulele care cuprind etichetele.

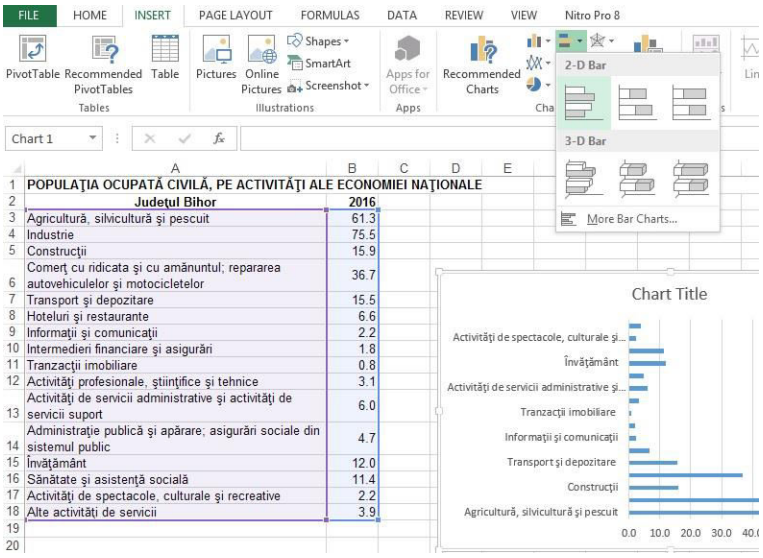
De exemplu: Dacă dorim să reprezentăm grafic populaţia ocupată civilă, pe activităţi ale economiei naţionale, înregistrată în anul 2016 (vezi tabelul de mai jos), nu este suficient să selectăm coloana cu numărul de persoane, ci vom selecta și categoriile (în cazul nostru tipurile de activități ale economiei).

Tabel: *Populaţia ocupată civilă, pe activităţi ale economiei naţionale, înregistrată în anul 2016*

Judeţul Bihor (mii de persoane)	**2016**
Agricultură, silvicultură şi pescuit	61.3
Industrie (total)	75.5
Construcţii	15.9
Comerţ cu ridicata şi cu amănuntul; repararea autovehiculelor şi motocicletelor	36.7
Transport şi depozitare	15.5
Hoteluri şi restaurante	6.6
Informaţii şi comunicaţii	2.2
Intermedieri financiare şi asigurări	1.8
Tranzacţii imobiliare	0.8
Activităţi profesionale, ştiinţifice şi tehnice	3.1
Activităţi de servicii administrative şi activităţi de servicii suport	6.0
Administraţie publică şi apărare; asigurări sociale din sistemul public	4.7
Învăţământ	12.0
Sănătate şi asistenţă socială	11.4
Activităţi de spectacole, culturale şi recreative	2.2
Alte activităţi de servicii	3.9

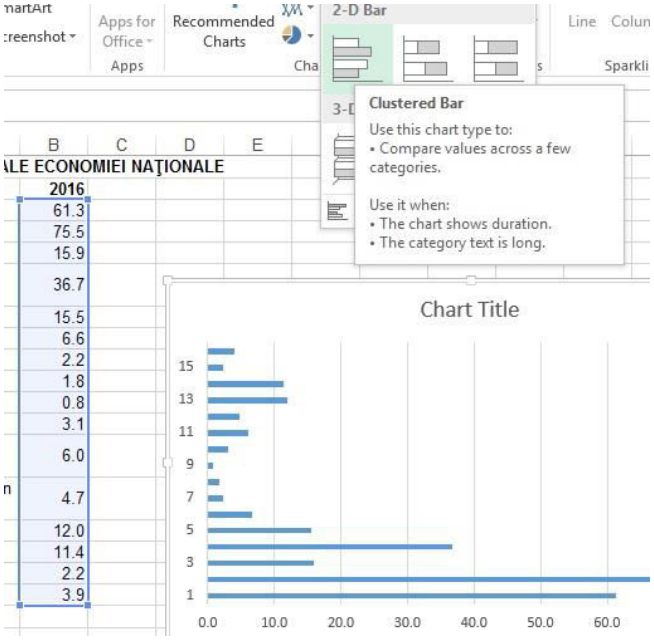
(Sursa datelor Direcţia Judeţeană de Statistică BIHOR, www.bihor.insse.ro)

Pentru a reprezenta grafic datele, selectăm ambele coloane, accesăm meniul **INSERT>Charts** și ne alegem tipul de grafic. În acest exemplu am selectat un grafic tip Bare (*Bar Chart*), dar într-o secțiune mai jos sunt prezentate detalii cu privire la alegerea tipurilor de grafice.



Graficul realizat cuprinde categoriile (în exemplul nostru tipuri de activități economice) și valorile înregistrate pe fiecare categorie. Puteți să observați că deocamdată nu se văd efectiv valorile (cifrele) aferente fiecărei categorii, dar urmează să le adăugăm ulterior.

Vom reveni asupra problemei privind selecția datelor. Dacă în exemplul nostru am selecta doar coloana reprezentând numerele, graficul obținut ar arăta astfel:

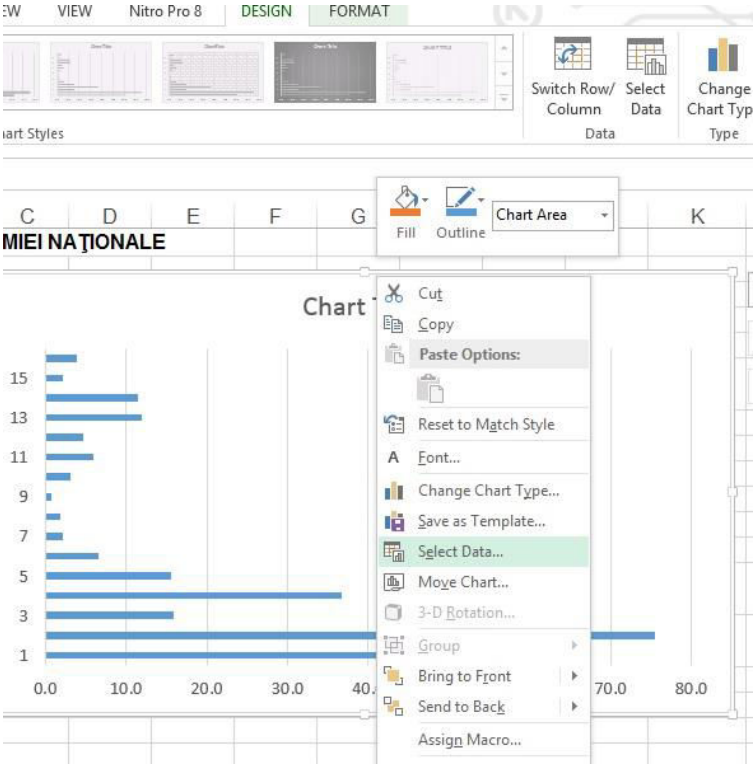


Graficul este corect, dar nu ne transmite informații complete, deoarece în locul categoriilor avem cifre: 15…13…11 etc.. Nu putem să citim informația cuprinsă în grafic, deoarece nu știm ce reprezintă aceste cifre. Trebuie să modificăm categoriile de pe axa respectivă.

Ce facem în situația în care graficul realizat nu cuprinde și valorile variabilelor?

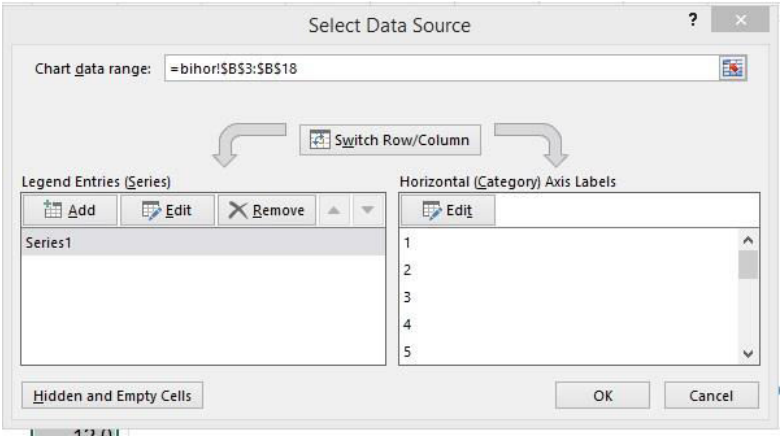
În această situație, avem două modalități de lucru: fie refacem graficul pe o selecție corectă, fie adăugăm categoriile.

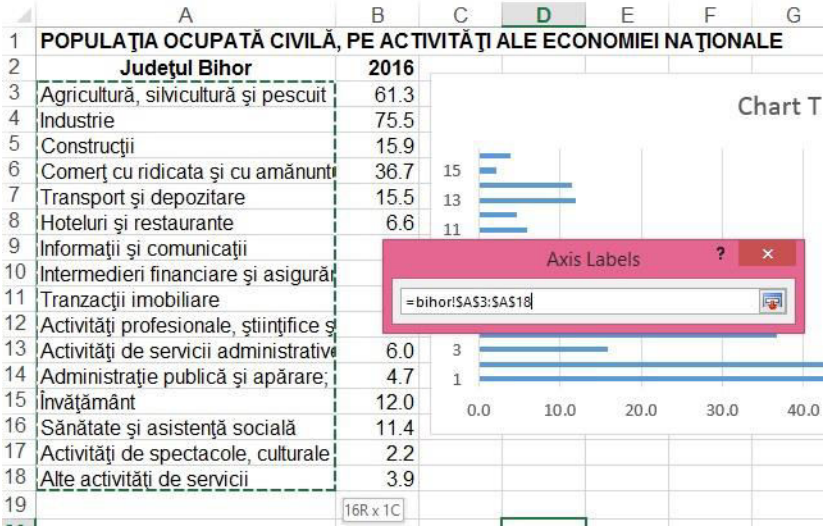
Pentru a adăuga categoriile pe axă, accesăm meniul **CHART TOOLS** (care este activ în momentul în care este selectat graficul) și vom alege **Select Data** (sau Click-dreapta pe grafic > **Select Data**).

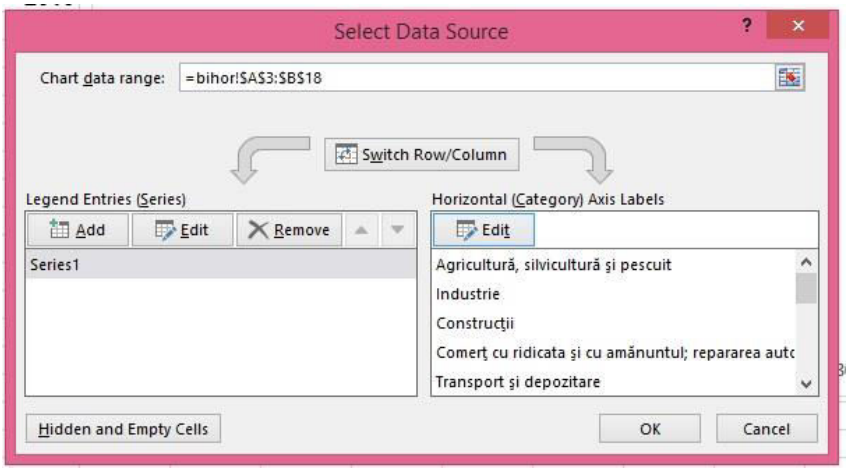


În fereastra privind sursa datelor (**Select Data Source**) care se deschide, în caseta din dreapta care cuprinde valorile categoriilor pe orizontală (*Horizontal Axis Labels*) trebuie să înlocuim etichetele prestabilite (cifrele) cu valorile corespunzătoare graficului nostru. Cu alte cuvinte, în exemplul nostru, trebuie să adaugăm aici coloana în care sunt specificate categoriile activităților economice. Pentru aceasta accesăm butonul **Edit**, apoi în fereastra nouă care se deschide (*Axis labels*) selectăm

direct celulele care conţin valorile din tabelul nostru (coloana unde sunt introduse categoriile activităţilor economice). Apoi confirmăm cu OK.





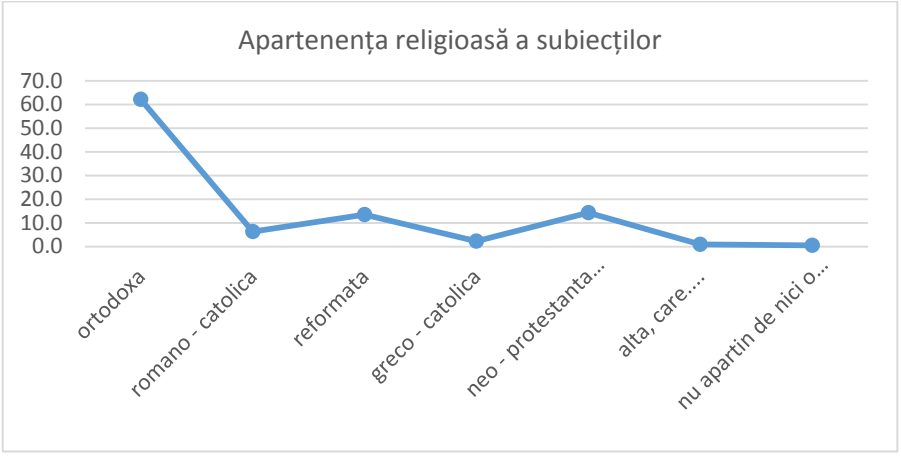


Cum alegem tipul de grafic?

Accesând meniul **INSERT>Charts**, avem o serie de tipuri de grafice din care putem să alegem. Alegerea va ține cont întotdeauna de tipul datelor pe care le avem la dispoziție. Putem să alegem între un grafic pe coloane (*Column Chart*), bare pe orizontală (*Bar Chart*), linie de reprezentare a evoluției (*Line*), suprafață (*Area*), grafic tip plăcintă (*Pie Chart*), nor de puncte (*Scatter*), radar (*Radar*) etc..

Cel mai des utilizate sunt graficele tip *Column*, *Bar* sau *Pie* Chart. Totuși, nu este recomandet să utilizăm un grafic tip plăcintă dacă datele includ mai mult de 8-10 categorii, deoarece "plăcinta" va fi împărțită în prea multe felii și va fi greu de urmărit informația. Nu este recomandat să alegem un grafic tip *Line*, dacă avem date măsurate la nivel nominal. Acest tip de grafic este potrivit pentru exemplificarea unei evoluții (dacă de exemplu am avea date înregistrate pe o anumită perioadă de timp).

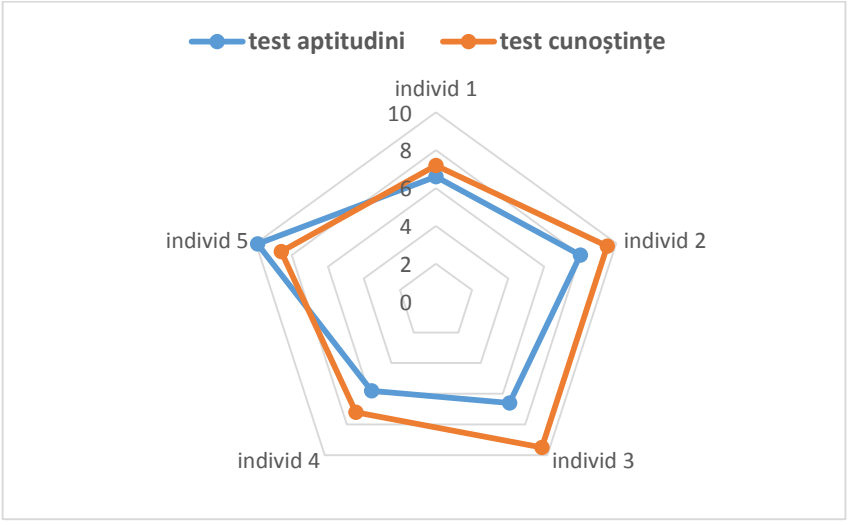
De exemplu: Dacă avem distribuția subiecților în funcție de religie, un grafic tip linie sugerează într-un fel o "scădere" de la religia ortodoxă la cea romano-catolică și o "creștere" spre cea reformată, ceea ce este greșit. O astfel de interpretare nu are sens.



Grafix: *Distribuția subiecților în funcție de religie*

Un grafic tip *Radial* este potrivit dacă dorim să exemplificăm îndepărtarea față de centru a anumitor categorii de indivizi.

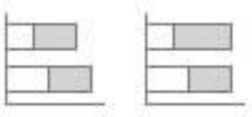
De exemplu: Dacă avem valorile obținute de către 5 indivizi în urma unui test de aptitudini și a unui test de evaluare a cunoștințelor, graficul tip *Radial* ne indică cine dintre cei 5 indivizi are rezultatele cele mai bune (Individul 5 este cel mai îndepărtat față de centru). Graficul ne indică faptul că toți indivizii, în afară de Individul 5, au obținut rezultate mai bune la testul de cunoștințe, comparativ cu rezultatele obținute la testul pentru aptitudini. Individul 3 obține rezultatele cele mai diferite între cele două teste.



Grafic: *Rezultatele obținute la cele două teste*

Fiecare tip de grafic cuprinde mai multe opțiuni privind prezentarea valorilor. Pe de o parte, putem alege între coloane 2D, respectiv 3D. În situația în care avem mai multe variabile în funcție de care sunt prezentate datele (acest tip de grafic presupune adăugarea legendei), putem să alegem ca aceste categorii să fie afișate în coloane alăturate sau să fie incluse într-o singură coloană împărțită pe mai multe culori (în funcție de numărul de categorii ale variabilei).

Ultimele două tipuri de grafice sunt asemănătoare. În această situație este important să alegem tipul de grafic corect, corespunzător datelor pe care dorim să le reprezentăm grafic. Vom alege ultimul tip doar dacă datele noastre sunt exprimate în procente. Acest tip de grafic ne ajută să comparăm valorile pe care le are fiecare procent raportat la total.



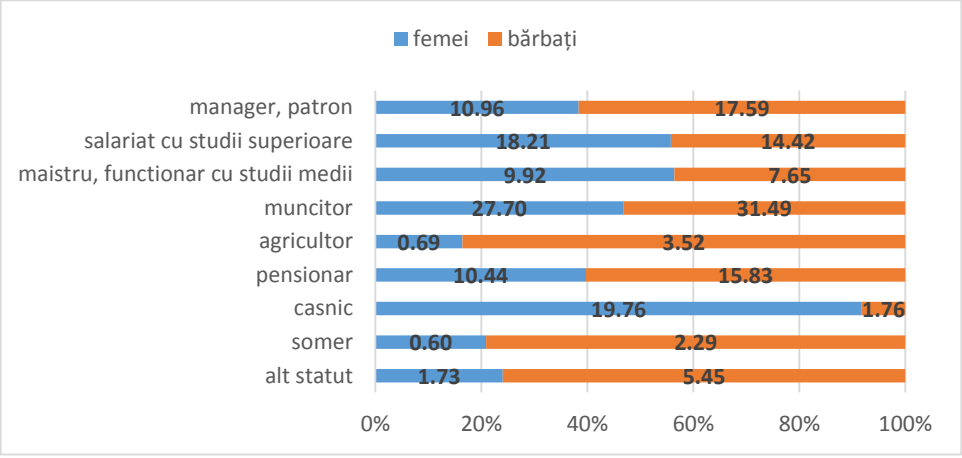
De exemplu: Dorim să reprezentăm grafic distribuția în funcție de statusul ocupațional, pe cele două genuri. Datele sunt înregistrate în procente, pe coloane. Vom citi astfel: dintre toate femeile înregistrate în eșantion, 10.96% sunt manageri, 18.21% sunt salariați cu studii superioare etc.. Totalurile se regăsesc pe coloane (total femei, respectiv total bărbați).

Tabel: *Distribuția în funcție de statusul ocupațional*

	femei	**bărbați**
manager, patron, director	10.96	17.59
salariat cu studii superioare	18.21	14.42
maistru, angajat cu studii medii	9.92	7.65
muncitor	27.70	31.49
agricultor	0.69	3.52
pensionar	10.44	15.83
casnic	19.76	1.76

şomer	0.60	2.29
alt statut	1.73	5.45

Putem să alegem tipul de grafic care prezintă datele în procente, dar important este să ne asigurăm că datele sunt reprezentate corect. În acest sens trebuie să ținem cont de modul în care sunt înregistrate datele și cum sunt raportate totalurile. Graficul de mai jos este unul greșit deoarece totalul procentelor nu este reprezentat corect.



Grafic: *Statusul ocupațional*

Putem să corectăm graficul prin înlocuirea datelor de pe coloane cu cele de pe rânduri. Pentru aceasta accesăm **CHART TOOLS > Switch Row/Column**.

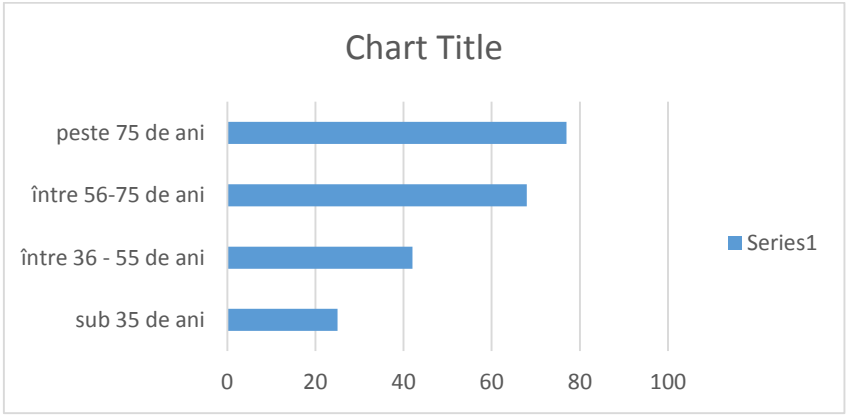
Cum ne asigurăm că graficul realizat cuprinde toate informațiile?

Pornind de la o selecție corectă a datelor, graficul este generat automat, dar acest lucru nu înseamnă că nu putem să adăugăm mai multe informații pentru a îmbunătăți aspectul graficului și pentru a înlesni modul de citire a datelor. Este important să verifică dacă graficul cuprinde toate informaţiile necesare, dacă informaţiile sunt corecte şi dacă sunt uşor de citit.

De exemplu: Să presupunem că avem următoarea distribuţie a numărului de beneficiari ai unor servicii sociale pe categorii de vârstă:

	A	B	C
1			
2		**vârsta**	**Numărul de beneficiari**
3		Sub 35 de ani	25
4		Între 36 şi 55 de ani	42
5		Între 56 şi 75 de ani	68
6		Peste 75 de ani	77
7			

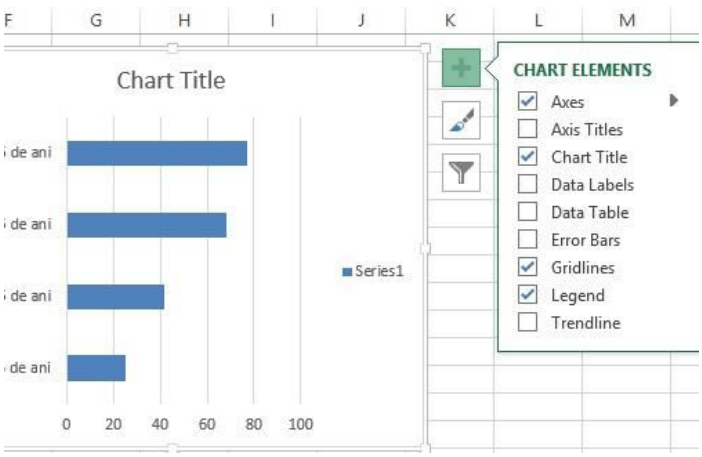
Pentru a realiza graficul accesăm meniul **INSERT>Charts** şi alegem tipul de grafic care se potriveşte pentru exemplificarea datelor noastre. Vom alege *Bar Chart*-ul.



Grafic: *Numărul de beneficiari ai serviciilor*

Pentru exemplul nostru privind beneficiarii de servicii sociale observăm că nu este necesară legenda (*Series1*), deoarece avem o singură coloană care reprezintă numărul de cazuri înregistrate pe o categorie de vârstă şi este folosită o singură culoare. Mai putem observa că numărul de cazuri înregistrate pe categoria de vârstă *Peste 75 de ani* sunt cele mai multe, dar nu ştim câte anume, deoarece graficul nu cuprinde şi valorile. Totodată, nu avem specificat titlul graficului.

Pentru a opera aceste modificări accesăm **Chart Elements** (primul buton marcat cu semnul **+**, din drepta graficului).

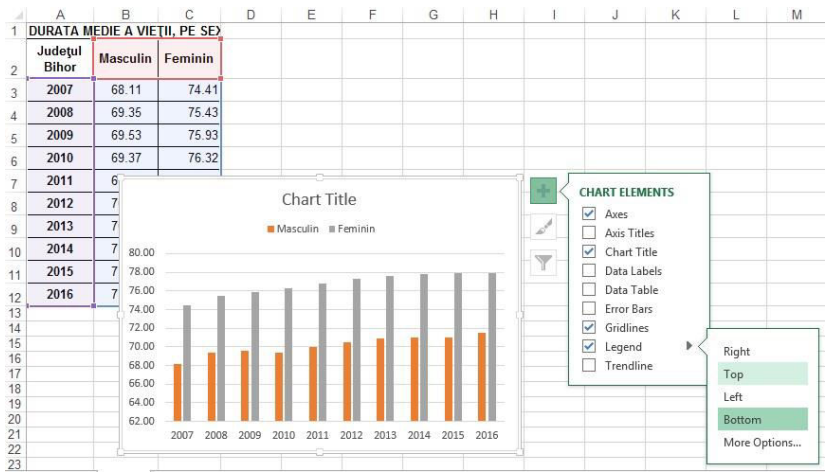


Aici avem lista tuturor elementelor: axele, titlul axelor, titlul graficului, valorile aferente etichetelor (*Data Labels*), tabelul pe baza căruia a fost realizat graficul (*Data Table*), legenda şi linia trendului (*Trendline*). Vom realiza modificările în concordanță cu tipul datelor/tipul graficului (de exemplu nu vom folosi legenda, dacă nu este cazul; vom adăuga întotdeauna valorile dacă spațiul permite etc.).

Sunt situații în care reprezentările grafice au nevoie de legendă. Aceasta se referă la specificarea categoriilor aferente culorilor incluse într-un grafic.

De exemplu: Dacă dorim să realizăm un grafic cu privire la durata medie a vieții pe cele două sexe, vom avea o culoare pentru valorile obținute în cazul bărbaților, respectiv o altă culoare pentru femei. Avem nevoie de o legendă care să ne indice care sunt categoriile pentru cele două culori.

Este important să pornim de la o selecție corectă, iar în acest caz, este nevoie să selectăm şi capul de tabel (primul rând din tabel care cuprinde valorile variabilei sex: feminin şi masculin). Putem să accesăm elementele pentru grafic (**Chart Elements**) şi să modificăm poziționarea legendei.

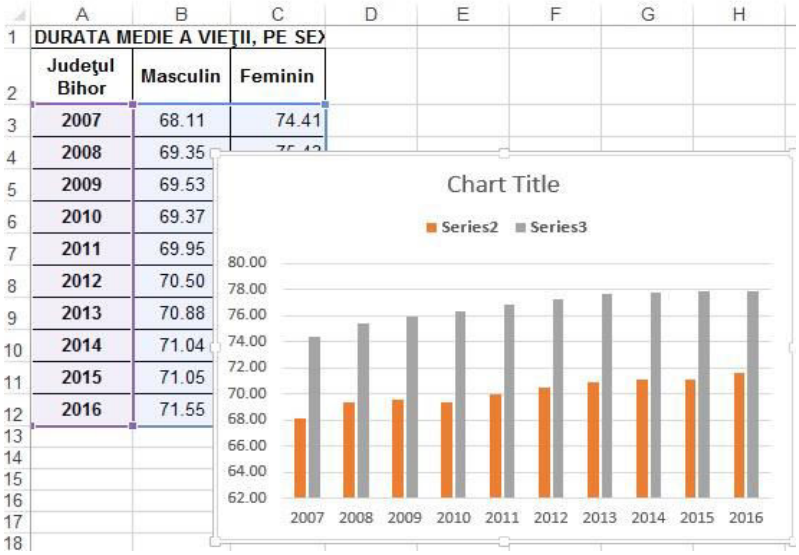


(sursa datelor: Direcţia Judeţeană de Statistică BIHOR, www.bihor.insse.ro)

Cum adăugăm informațiile pentru legendă, dacă acestea lipsesc?

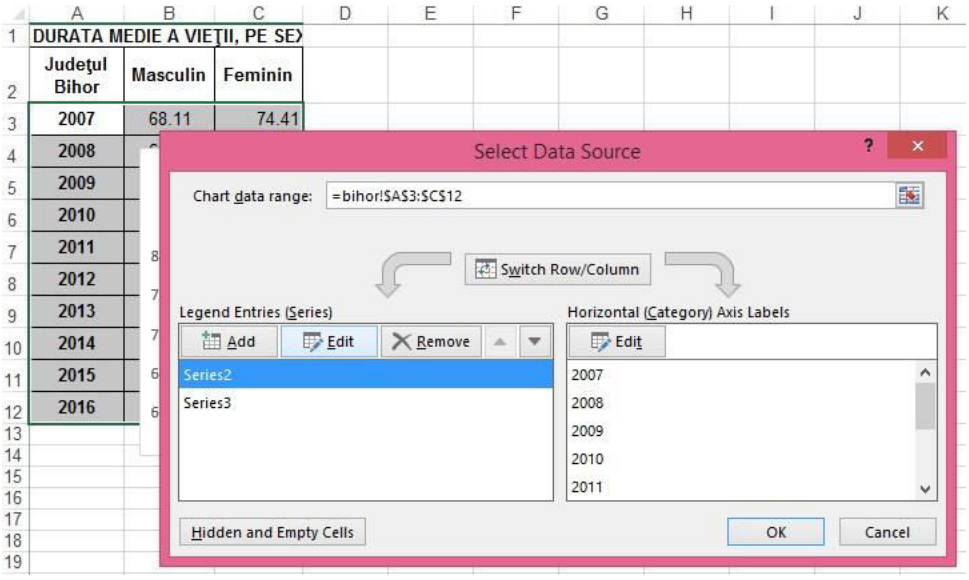
În cazul în care graficul necesită includerea unei legedne, iar aceasta nu este afișată corect (pornind de la o selecție a datelor incorectă), este necesar să adăugăm ulterior categoriile aferente legendei.

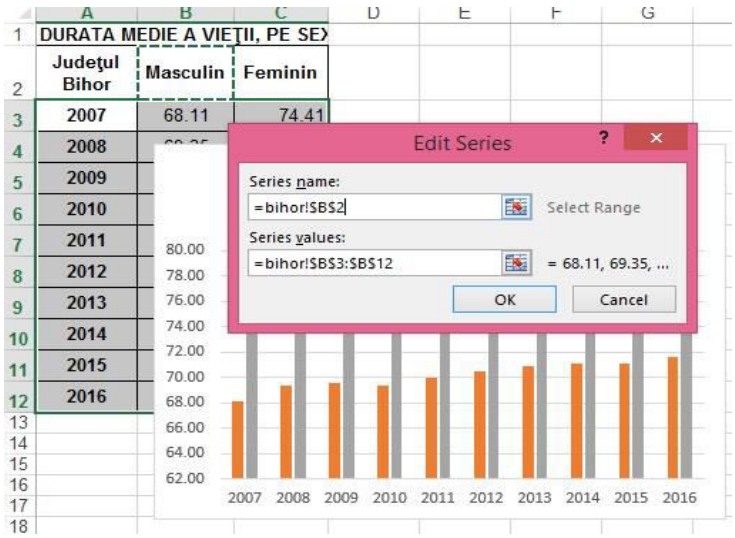
De exemplu: Dacă pentru realizarea graficului privind durata medie a vieții pe cele două sexe, am fi pornit doar de la selecția numerelor din coloane, fără să includem capul de tabel (unde sunt specificate categoriile Masculin și Feminin), graficul ar arăta astfel:



Este dificil să citim informația inclusă în grafic, deoarece nu știm ce reprezintă *Series2*, respectiv *Series3*. În locul acestora trebuie să adăugăm categoriile specifice datelor noastre (Masculin și Feminin).

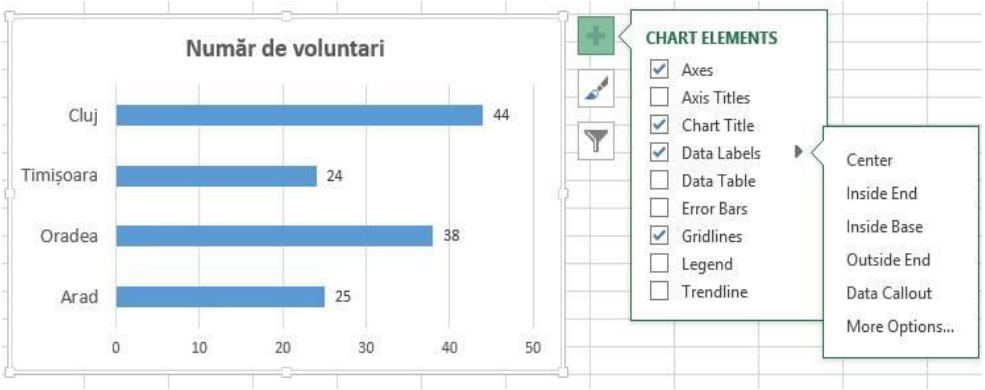
Accesăm meniul **CHART TOOLS > Select Data** (sau Click-dreapta pe grafic > **Select Data**). În caseta din stânga avem intrările pentru legendă (**Legend Entries).** Selectăm fiecare serie în parte, accesăm butonul **Edit** și selectăm din tabel categoria corespunzătoare seriei. În exemplul nostru categoria Masculin se regăsește în celula B2, respectiv pentru Feminin vom selecta celula C2.



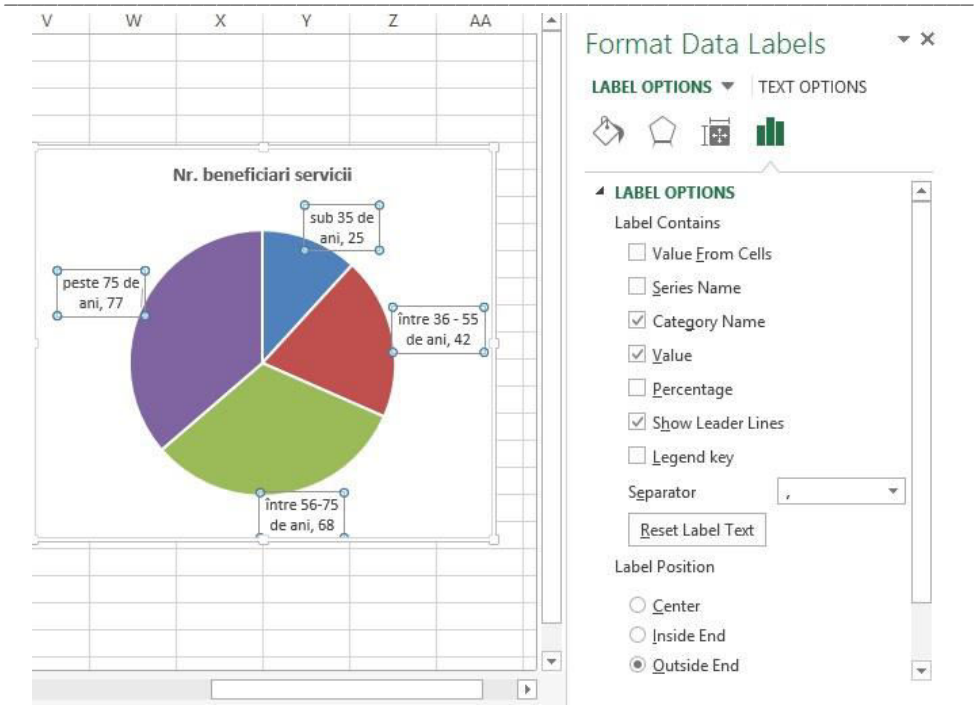


Cum adăugăm valorile etichetelor?

În ceea ce privește adăugarea valorilor pentru categoriile (etichetele) incluse în grafic (***Data Labels***), este recomandat să le adăugăm în toate graficele, deoarece asigură afișarea informațiilor complete și nu mai este necesar să revenim la tabelul care cuprinde datele. Sunt totuși situații în care valorile adăugate ar încărca prea mult graficul. În acest caz, putem să adăugăm tabelul direct în grafic (***Data Table***).

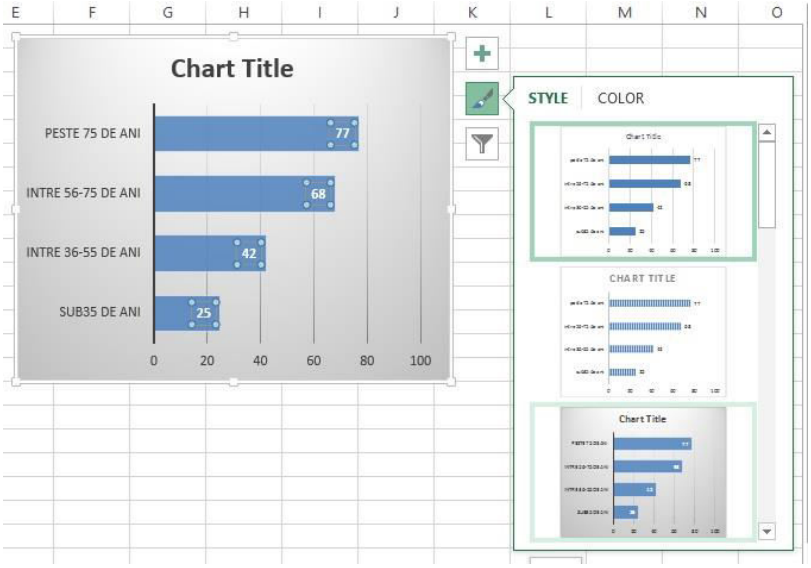


Pentru grafice tip *Pie Chart* putem să optăm să fie afișată atât valoarea (***Value***), cât și categoria (***Category Name***).

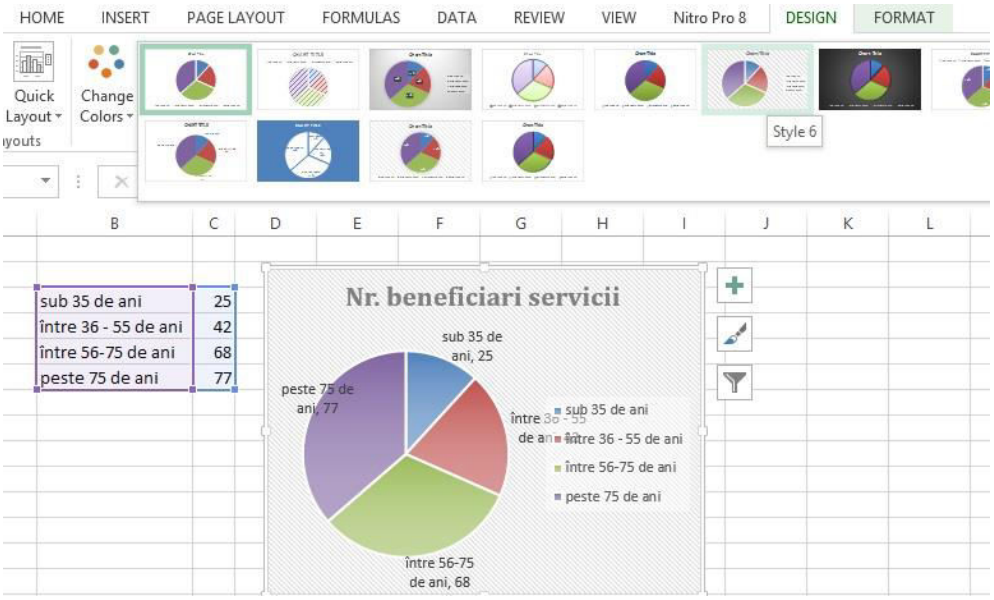


Cum modificăm design-ul graficului?

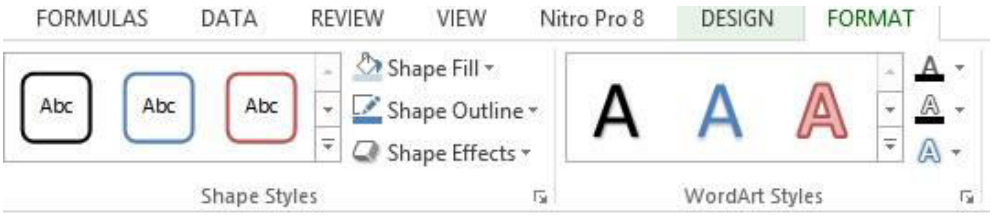
Accesarea celui de-al doilea buton (în formă de pensulă, din dreapta graficului) permite utilizarea diferitelor stiluri pentru grafic sau modificarea culorilor.



Putem să facem modificări privind aspectul graficului accesând meniurile **DESIGN**, respectiv **FORMAT**, care sunt active în momenul în care graficul este selectat.

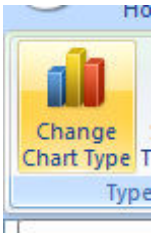


Accesând meniul **FORMAT**, putem să facem modificări ale obiectului selectat (culoarea din interior, culoarea liniei, efecte de umplere) și modificări ale textului (culoarea textului, efecte de umplere etc.).



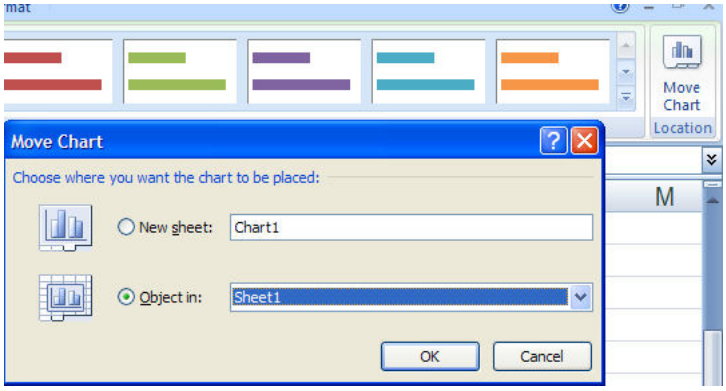
Cum modificăm un grafic?

Dacă dorim să modificăm tipul de grafic utilizat accesăm din meniul **Chart Tools>DESIGN>Change Chart Type.** Această opțiune ne deschide direct fereastra care cuprinda lista tipurilor de grafice.



Cum mutăm un grafic?

Pentru a modifica locația graficului accesăm **Chart Tools>DESIGN>Move Chart.** În fereastra care se deschide putem să alegem să mutăm graficul fie ca obiect în altă pagină de lucru, fie pe o pagină separată.

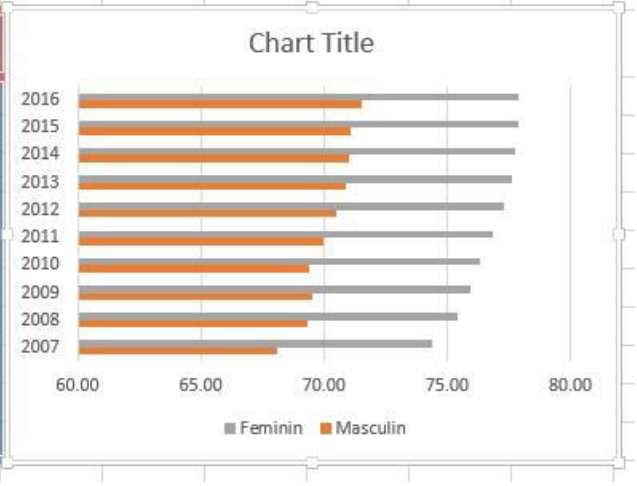


Cum realizăm formatări suplimentare ale graficului?

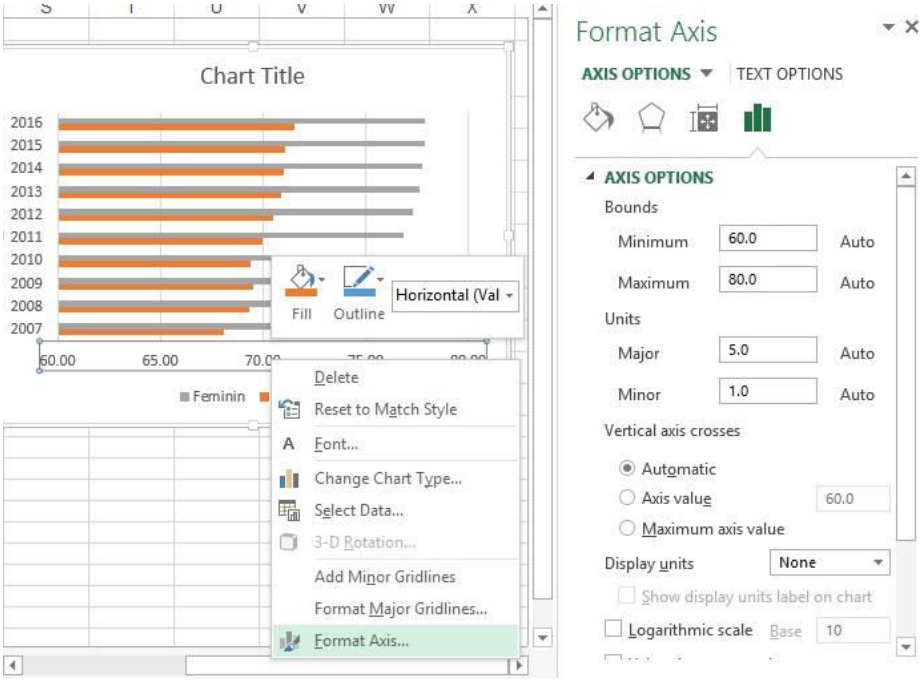
După realizarea unui grafic, în funcție de modul în care sunt afișate informațiile, putem să facem modificări suplimentare în ceea ce privește spațiul dintre axe (*Format Plot Area*), suprafața graficului (*Format Chart Area*), suprafața aferentă seriei de date (*Format Data Series*), axele (*Format Axis*). Aceste modificări le facem accesând Click-dreapta > ***Format.*** Obiectul pe care l-am selectat, acela va fi modificat.

De exemplu: În reprezentarea grafică a datelor privind speranța medie cu viața, pe cele două genuri, axa pornește de la vârsta de 60 de ani și este marcată din 5 în 5 ani.

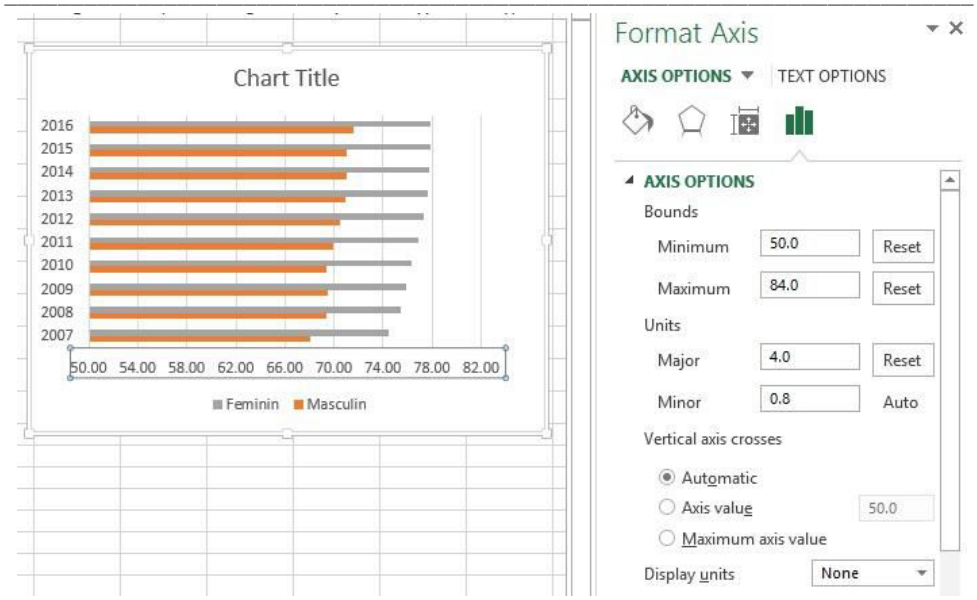
	A	B	C
1	DURATA MEDIE A VIEȚII, PE SEX		
2	Județul Bihor	Masculin	Feminin
3	2007	68.11	74.41
4	2008	69.35	75.43
5	2009	69.53	75.93
6	2010	69.37	76.32
7	2011	69.95	76.84
8	2012	70.50	77.28
9	2013	70.88	77.63
10	2014	71.04	77.76
11	2015	71.05	77.89
12	2016	71.55	77.9
13			



Dacă dorim să modificăm axa, accesăm *Click-dreapta>Format Axis.*



Aici putem să modificăm valoarea minimă, respectiv maximă a axei și să schimbăm unitatea de măsură (în loc de 5 ani, putem să selectăm 4 ani).



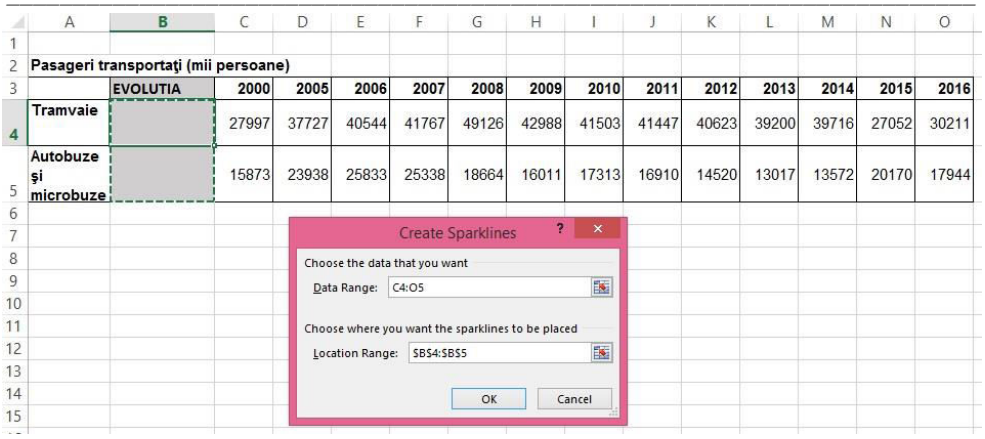
Alte modalități de prezentare vizuală a datelor

În funcție de cantitatea datelor şi tipul acestora, sunt situații în care putem opta să nu realizăm un grafic pe baza datelor pe care le avem la dispoziție, ci să realizăm diagrame direct în celule. Acestea se utilizează în mod special când dorim să evidențiem evoluția unei serii de date. Utilizarea **Sparkline** asigură prezentarea vizuală a tendinței evoluției datelor.

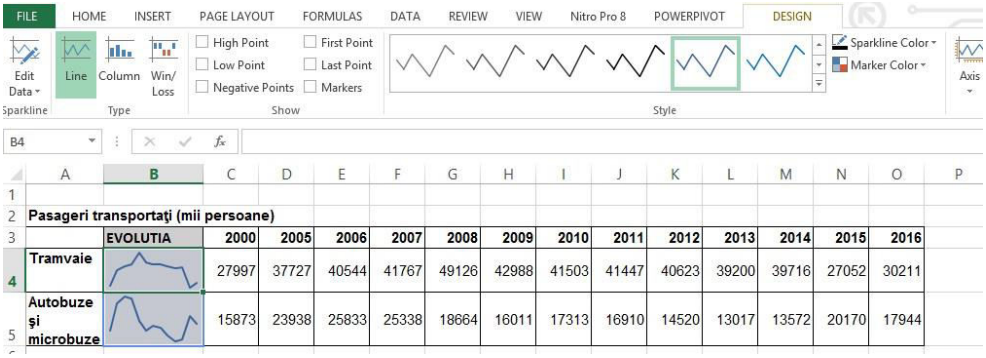
De exemplu: Avem înregistrat numărul de pasageri pentru transportul urban din județul Bihor în perioada 2000 – 2016 (sursa datelor: Directia Judeteana de Statistica Bihor). Pentru a vizualiza trendul înregistrat în perioada analizată, putem să adăugăm o coloană care va trasa liniile trendului.

Accesăm **INSERT > Sparklines>Line.**

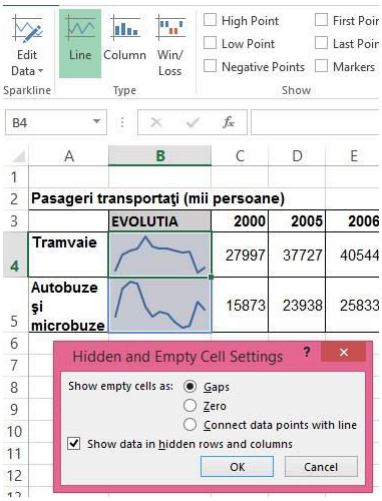
Pentru ***Data Range*** vom selecta seria de date, iar apoi vom adăuga o locație (selecția de celule) în care să fie introduse liniile. În cazul nostru vom selecta coloana nouă inserată în care vom vedea EVOLUȚIA.



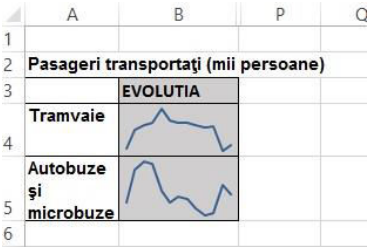
Selectarea celulelor în care am adăugat ***Sparkline*** ne deschide meniul unde găsim o serie de opțiuni pentru modificarea modului de vizualizare a liniilor (culoare, grosime adăugare marcatori etc.).



Dacă dorim să ascundem (*Hide*) coloanele care conțin datele și să păstrăm doar liniile, vom selecta din meniul **Sparkline > Edit Data > Hidden and Empty Cells**… *Show data in hidden rows and columns*.



Vom ascunde coloanele iar fişa noastră de lucru va arăta astfel:



Câteva recomandări și observații finale privind reprezentarea grafică

Nu este recomandat să realizăm un grafic doar pentru a ocupa cât mai mult spațiu într-un raport sau o lucrare.

Este important să nu încărcăm excesiv graficul (culori, tipuri de caractere ect.). Ceea ce ne interesază este informația prezentată în grafic şi mai puțin aspectul graficului în sine.

Este bine să evităm informațiile redundante: fie realizăm un grafic care să cuprindă toate informațiile, fie adăugăm datele într-un tabel. Nu adăugăm şi un grafic şi un tabel conținând aceleaşi date.

Pentru exemplificarea grafică corectă a datelor este foarte important să ţinem cont de modul în care vor fi citite datele: includem toate informaţiile, nu doar prezentări parţiale; ne asigurăm că informaţiile sunt corecte; alegem tipul de grafic care se potriveşte pentru tipul datelor disponibile.

9 UTILIZAREA FUNCȚIILOR PENTRU GESTIUNEA INFORMAȚIILOR

CONCATENATE, PROPER, MID, LEFT, RIGHT, LEN, UPPER, LOWER, ISEVEN, ISODD, ISNUMBER, ISTEXT

Programul Excel este utilizat în principal pentru prelucrarea şi analiza datelor. Sunt însă situații în care este nevoie să utilizăm anumite funcții pentru managementul informațiilor, în funcție de tipul acestora. Aceste funcții se regăsesc în categoria funcțiilor pentru text şi cea a funcțiilor pentru informații.

Funcțiile pentru text

Funcția
= CONCATENATE(text1, [text2],…)

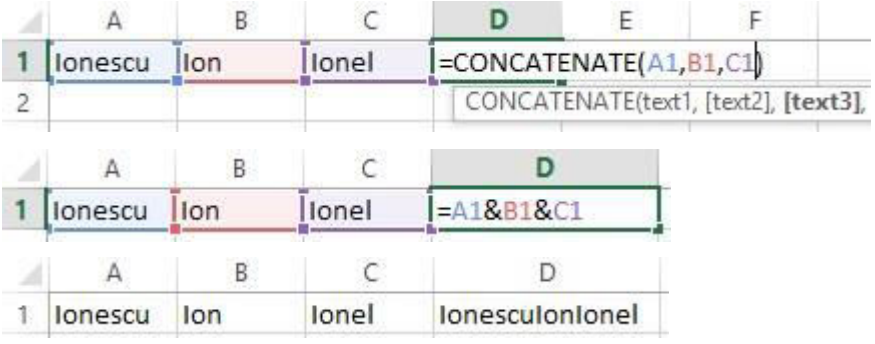
Descriere

Funcția unește caracterele din mai multe celule între-un singur şir. Cu funcția **=CONCATENATE** putem adăuga caractere în plus față de caracterele din celulele selectate. Pentru a uni un şir de caractere putem folosi şi semnul & între referințe.

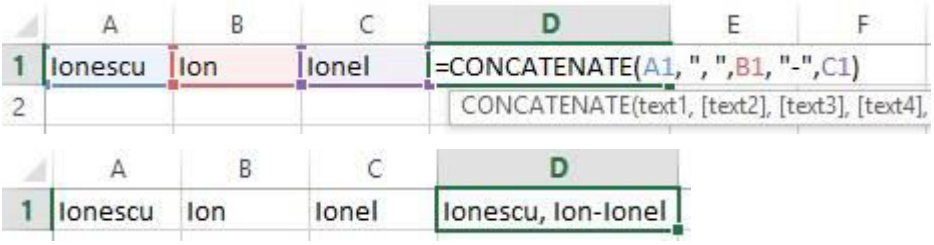
Argumentele

Text1, text 2 – referințele care conțin caracterele pe care dorim să le unim.

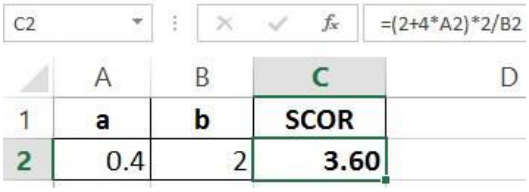
De exemplu: Dacă avem numele şi prenumele introduse în celule separate şi dorim informația să apară într-o singură celulă, vom folosi funcția **=CONCATENATE**.


A B C D E F
1 Ionescu Ion Ionel =CONCATENATE(A1,B1,C1)
2
CONCATENATE(text1, [text2], [text3],
A B C D
1 Ionescu Ion Ionel =A1&B1&C1
A B C D
1 Ionescu Ion Ionel IonescuIonIonel

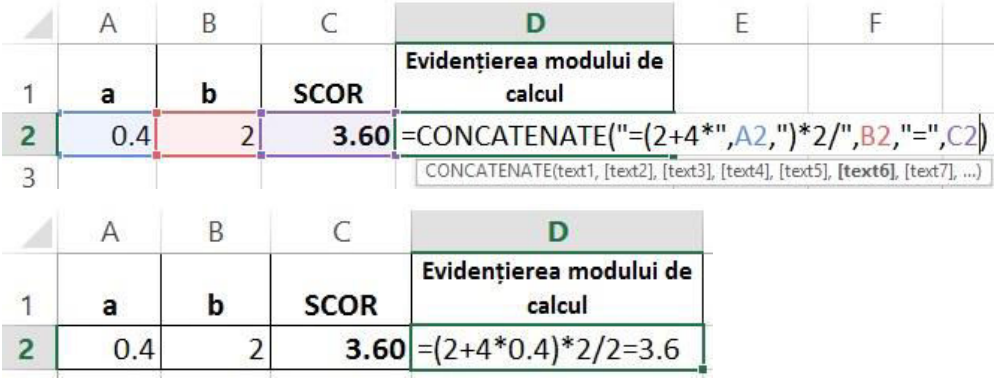
Putem să adăugăm între argumentele funcției anumite caractere pentru a schimba formatul:


A B C D E F
1 Ionescu Ion Ionel =CONCATENATE(A1, ", ",B1, "-",C1)
2
CONCATENATE(text1, [text2], [text3], [text4],
A B C D
1 Ionescu Ion Ionel Ionescu, Ion-Ionel

Un alt exemplu: Avem în foaia de lucru un scor calculat pe baza unei formule în care sunt incluse două valori: a și b. Formula utilizată este: =(2+4*a)*2/b.



Dacă dorim să evidențiem modul în care a fost calculat scorul respectiv, vom folosi funcția **=CONCATENATE**, în care putem adăuga caractere pentru a oferi detalii cu privire la modul de calcul.


A B C D E F
1 a b SCOR Evidențierea modului de calcul
2 0.4 2 3.60 =CONCATENATE("=(2+4*",A2,")*2/",B2,"=",C2)
3
CONCATENATE(text1, [text2], [text3], [text4], [text5], [text6], [text7], ...)
A B C D
1 a b SCOR Evidențierea modului de calcul
2 0.4 2 3.60 =(2+4*0.4)*2/2=3.6

Funcția

= PROPER(text)

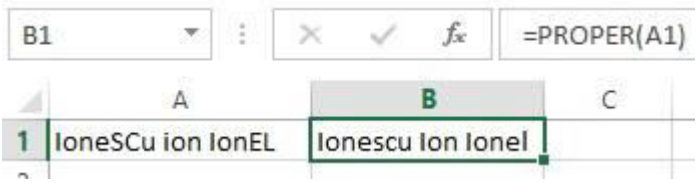
Descriere

Funcția convertește caracterele (textul) dintr-o celulă având ca format: primul caracter cu literă mare, iar restul caracterelor cu litere mici .

Argumentele

Text – caracterele sau referința pe care dorim să le convertim.

De exemplu:



Funcțiile

= LOWER(text)

= UPPER(text)

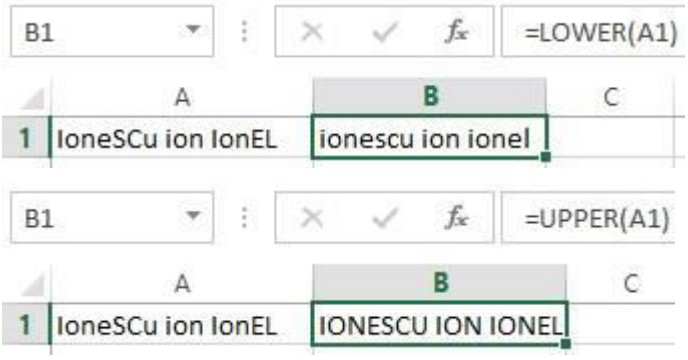
Descriere

Funcțiile convertesc caracterele dintr-o referință în text cu litere mici (LOWER), respectiv text cu litere mari (UPPER.

Argumentele

Text – caracterele sau referința pe care dorim să le convertim.

De exemplu:



Funcția

= LEN(text)

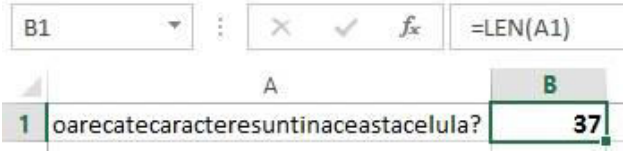
Descriere

Funcția returnează numărul de caractere dintr-un șir.

Argumentele

Text – caracterele pe care dorim să le numărăm.

De exemplu:



Funcțiile

= LEFT(text, [num_chars])

= RIGHT(text, [num_chars])

Descriere

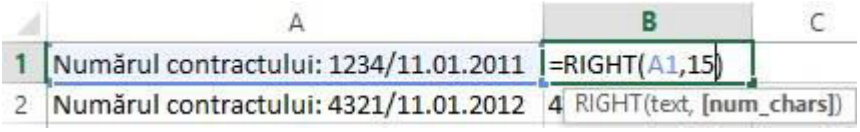
Funcțiile extrag dintr-un șir de caractere un număr specificat de caractere de la stânga, respectiv de la dreapta.

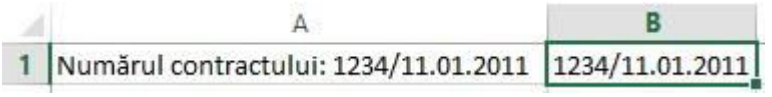
Argumentele

Text – șirul de caractere din care dorim să extragem, de la stânga, sau de la dreapta, un anumit număr de caractere.

Num_chars – numărul de caractere pe care dorim să le extragem.

De exemplu:





Funcția

= MID(text, start_num, num_chars)

Descriere

Funcția extrage din mijlocul unui șir de caractere un anumit număr specificat de caractere.

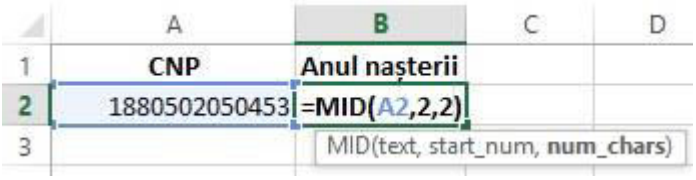
Argumentele

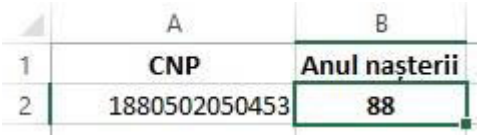
Text – șirul de caractere

Start_num – poziția caracterului de la care să înceapă extragerea.

Num_chars – numărul de caractere extrase.

De exemplu:





Funcțiile pentru informații

Funcția

= ISTEXT(value)

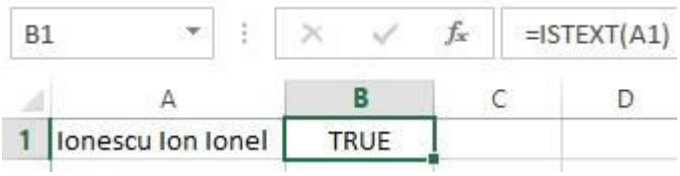
Descriere

Funcția verifică dacă informația dintr-o celulă este de tip text și returnează rezultatul TRUE (ADEVĂRAT) dacă informația este de tip text, respectiv FALSE (FALS) dacă are alt format.

Argumentele

value – celula în care se află informația ce trebuie verificată.

De exemplu:



Funcția

= ISNUMBER(value)

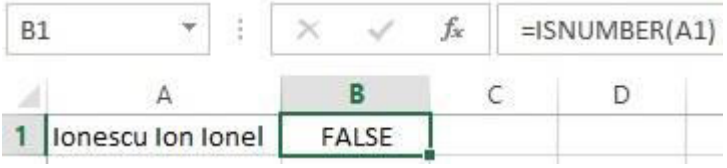
Descriere

Funcția verifică dacă informația dintr-o celulă este de tip număr și returnează rezultatul TRUE (ADEVĂRAT) dacă informația este de acest tip, respectiv FALSE (FALS) dacă are alt format.

Argumentele

value – celula în care se află informația ce trebuie verificată.

De exemplu:



Funcțiile

= ISEVEN(number)

=ISODD(number)

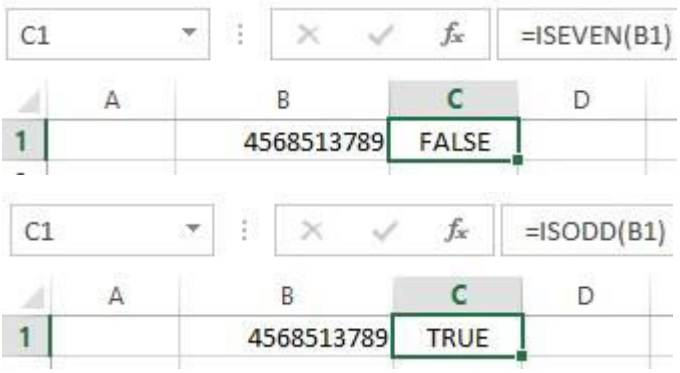
Descriere

Funcțiile verifică dacă numărul dintr-o celulă este par (EVEN), respectiv impar (ODD), și returnează rezultatul TRUE (ADEVĂRAT) dacă numărul este de acest tip, respectiv FALSE (FALS) dacă nu este de acest tip.

Argumentele

number – celula în care se află numărul ce trebuie verificat.

De exemplu:



Resurse bibliografice recomandate

Carlberg, C. (2014). Statistical Analysis: Microsoft Excel 2013, US: Que Publishing.

Coman, C. (2011). Statistica aplicată în științele sociale, București: Institutul European.

Constantinescu R. & Danaila I. (2017). Manual ECDL Calcul tabelar - Microsoft Excel 2016, București: Editura ECDL ROMANIA.

Frye, C.D. (2013). Microsoft Excel 2013. Step by Step, California: O'Reilly Media, Inc.

McFedries, P. (2013). Excel 2013 Formulas and Functions, US: Que Publishing.

Moriarty, B., Held, B. și Richardson, T. (2018). Microsoft Excel Functions & Formulas, US: Stylus Publishing, LLC.

Rotariu, T., Badescu G., Culic, I. Mezei, E. și Muresan, C. (1999). Metode statistice aplicate în științele sociale. Iași: Editura Polirom.

ISBN: 978-606-37-0822-0